Benjamin McKay

Stokes's Theorem

May 21, 2013

This work is licensed under a Creative Commons Attribution-ShareAlike 3.0 Unported License.

Preface

This course proves Stokes's theorem, starting from a background of rigorous calculus. Chapters 9 and 10 of Rudin [6] cover the same ground we will, as does Spivak [7] and Hubbard and Hubbard [4]. I encourage you to read those books. All of them, including these notes, are attempts to rewrite Milnor's little book [5], and try to lead up to Bott and Tu [2] and Guillemin and Pollack [3].

There is one abstract idea in this book: differential forms. Let's consider 5 motivations for pursuing this abstract idea.

1. Look at an integral, like
$$\int x^2\,dx.$$
What is dx? Recall
$$\int f(x)\,dx = \lim_{\Delta x \to 0} \sum f\left(x_i\right) \Delta x,$$
so dx plays the role of Δx. Since $\Delta x \to 0$, we think of dx as being "infinitesimal" (infinitely small). But that is nonsense. Physicists play with infinitesimals as if they made sense. We will give dx a meaning in this book, so that we can play with the physicists. This dx is the simplest example of a differential form.
2. We want to carry out integrals over geometric objects: curves and surfaces. To calculate those integrals we have to parameterize those objects, by mapping pieces of the real number line or the plane to the curves or surfaces. The integral will make sense once we check that the result doesn't depend on how we parameterize. Differential forms yield reparameterization invariant integrals.
3. Many integrals depend on the direction we integrate. For example, in single variable calculus, $\int_1^0$ means $-\int_0^1$, which is the only definition that makes all of the theory hold identically no matter which direction you integrate. Similar phenomena occur in higher dimensions; we need to keep track of signs. The notation of differential forms keeps track of the signs for us. For multidimensional integrals, this involves a little algebra.
4. Differential forms yield a bridge between problems in topology and problems in calculus. We will use them to prove Brouwer's fixed point theorem.
5. Differential forms arise naturally in the theory of electromagnetic fields; we won't pursue this direction.

Contents

1 Euclidean Space 1

2 Differentiation 11

3 The Contraction Mapping Principle 21

4 The Inverse Function Theorem 29

5 The Implicit Function Theorem 33

6 Manifolds 39

7 Integration 49

8 Vector Fields 53

9 Differential Forms 57

10 Differentiating Differential Forms 63

11 Integrating Differential Forms 69

12 Stokes's Theorem 77

13 The Brouwer Fixed Point Theorem 85

14 Manifolds from Inside 91

15 Volumes of Manifolds 97

16 Sard's Theorem 105

17 Homotopy and Degree 111

Hints 121

Bibliography 133

List of Notation 135

Index 137

Chapter 1

Euclidean Space

We recall definitions from previous courses and discuss compactness. Closed bounded sets of points are called compact. Compactness has some subtle consequences.

Maps

We use the usual terminology and notation of sets without introduction. We write $\mathbb{R}$ to mean the set of all real numbers. A *map* or *function* $f \colon X \to Y$ is a rule associating to any point x in some set X a point y in the set Y. We assume the reader is familiar with composition of functions, inverse functions, and what it means to say that a function is 1-1 (also known as injective), or is onto (also known as surjective).

Suppose that $f \colon X \to Y$ is a map between sets and $S \subset X$ is a subset. The *image* $f(S)$ of S is the set of all points $f(x)$ for all $x \in S$. The *image* of f is $f(X)$. Similarly if $T \subset Y$ is a subset, the *preimage*, $f^{-1}T$, of T is the set of points $x \in X$ for which $f(x) \in T$. It will often be convenient to avoid choosing a name for a function, for example writing $x \mapsto x^2 \sin x$. to mean the function

$$f \colon \mathbb{R} \to \mathbb{R}, f(x) = x^2 \sin x.$$

Euclidean space

The set $\mathbb{R}^n$ is the set of all n-tuples

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

of real numbers $x_1, x_2, \dots, x_n \in \mathbb{R}$. Following standard practice, we will often be lazy and write such a tuple horizontally as

$$x = (x_1, x_2, \dots, x_n) .$$

1.1 For X and Y subsets of points in the following sets, how might we try to draw a picture to describe a map $f \colon X \to Y$?

(a) $X \subset \mathbb{R}, Y \subset \mathbb{R}$,
(b) $X \subset \mathbb{R}^2, Y \subset \mathbb{R}$,

(c) $X \subset \mathbb{R}, Y \subset \mathbb{R}^2$,
(d) $X \subset \mathbb{R}^3, Y \subset \mathbb{R}$,
(e) $X \subset \mathbb{R}, Y \subset \mathbb{R}^3$,
(f) $X \subset \mathbb{R}^2, Y \subset \mathbb{R}^2$

The vectors in $\mathbb{R}^n$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

are called the *standard basis vectors*. The *norm* of a vector $x \in \mathbb{R}^n$ is

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

The *inner product* of two vectors $x, y \in \mathbb{R}^n$ is

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

The *distance* between two points $x, y \in \mathbb{R}^n$ is $\|x - y\|$.

1.2 Prove the triangle inequality: for any points $p, q, r \in \mathbb{R}^n$,

$$\|p - q\| + \|q - r\| \geq \|p - r\|.$$

The *ball* or *open ball* of radius r around a point $x \in \mathbb{R}^n$ is the set $B_r(x)$ of all points of $\mathbb{R}^n$ of distance less than r from x. The *closed ball* of radius r around a point $x \in \mathbb{R}^n$ is the set $\bar{B}_r(x)$ of all points $p \in \mathbb{R}^n$ of distance less than or equal to r from x. A set is *bounded* if it lies in a ball. A map is *bounded* if its image is bounded.

For any real numbers $a, b \in \mathbb{R}$, we write $[a, b] \subset \mathbb{R}$ to mean the set of points $x \in \mathbb{R}$ so that $a \leq x \leq b$, and we call $[a, b]$ the *closed interval* from a to b. A *box* in $\mathbb{R}^n$ is a subset of the form

$$X = [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_n, b_n],$$

i.e. a product of closed intervals. In other words, a vector $x \in \mathbb{R}^n$ lies in the box X just when $a_1 \leq x_1 \leq b_1$ and $a_2 \leq x_2 \leq b_2$ and …and $a_n \leq x_n \leq b_n$.

1.3 Prove that every ball of radius r lies inside a box all of whose sides has length at most $2r$. Prove that every bounded set lies inside a box.

1.4 Prove that every box is bounded. Be precise: if all sides have length no more than δ, prove that your box lies in an open ball of any radius larger than $\delta\sqrt{n}/2$.

Open sets

A set $U \subset \mathbb{R}^n$ is *open* if around every point $x \in U$ there is some ball $B_r(x)$ of positive radius contained entirely in U. Intuitively, x is “surrounded by U”, or in other words lies deep inside U, not on the surface. Intuitively, U must be a fat blob, not a skinny thing like a line or a curve or an isolated point. This is why an open set containing a point is often called an open set “around” that point. A set is called *closed* when its complement is open.

1.5* Prove that every box is closed.

1.6 Prove that the half open interval I consisting of the numbers x with $0 \le x < 1$ is not open *and* also not closed.

A set is a *neighborhood* of a point if that set contains an open set around that point.

We will often write

$$f\colon \text{open } \subset \mathbb{R}^q \to \mathbb{R}^p$$

to mean that f is a map taking some unspecified open subset of $\mathbb{R}^q$ to $\mathbb{R}^p$. A map is *continuous* if the preimage of any open set is open.

1.7 Prove that a map $f\colon X \to Y$ between subsets $X \subset \mathbb{R}^q$ and $Y \subset \mathbb{R}^p$ is continuous just when, for every point $x_0 \in X$, and for any positive number ε, there is a positive number δ so that for every point x closer than δ to x_0, $f(x)$ is closer than ε to $f(x_0)$.

It follows easily that all arithmetic operations are continuous where defined, and that compositions of continuous maps are continuous.

1.8 Prove that any continuous integer-valued function of a real variable is constant.

1.9 Find a continuous map f and an open set U so that $f(U)$ is not open.

1.10 Find a continuous map f and an closed set A so that $f(A)$ is not closed.

1.11* Prove that the map taking a square matrix to its inverse is continuous where defined.

1.12 Denote by $\operatorname{GL}(n, \mathbb{R})$ the set of invertible $n \times n$ matrices. Prove that $\operatorname{GL}(n, \mathbb{R}) \subset \mathbb{R}^{n^2}$ is an open set.

If $A \subset \mathbb{R}^n$ is any subset, a subset $B \subset A$ is *open subset* of A if $B = A \cap U$ for some open subset $U \subset \mathbb{R}^n$. Similarly, we define continuity of maps between subsets, etc.

Convergence

A sequence $x_1, x_2, \dots \in \mathbb{R}^n$ *converges* to a point $x \in \mathbb{R}^n$ just when every open set containing x contains all but finitely many of the points in the sequence. We write

$$x = \lim_{j \to \infty} x_j \text{ or } x_j \to x$$

and say that x is the *limit* of the sequence. Similarly, if $X \subset \mathbb{R}^q$ and $Y \subset \mathbb{R}^p$ are sets and $f \colon X \to Y$ is a map, and $x_0 \in X$, write

$$\lim_{x \to x_0} f(x) = y_0$$

to mean that the preimage of every neighborhood of y_0 is a neighborhood of x_0.

1.13 Restate the definitions of limit in terms of ε and δ.

1.14 Prove that a sequence $x_1, x_2, x_3, \dots \in \mathbb{R}^n$ converges to x just when each of the entries in the vectors

$$x_1 = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1n} \end{pmatrix}, x_2 = \begin{pmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{pmatrix}, \dots$$

converges to the corresponding entry in

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

1.15 Prove that a sequence $x_1, x_2, x_3, \dots \in \mathbb{R}^n$ converges to x just when the map

$$f \colon \{1, 2, 3, \dots\} \to \mathbb{R}^n, f(j) = x_j$$

is continuous.

A sequence $x_1, x_2, \dots \in \mathbb{R}^n$ is a *Cauchy sequence* if, no matter how close I want the elements of the sequence to be to one another, all but finitely many of them are at least that close to one another.

1.16 An *upper bound* for a set of numbers $S \subset \mathbb{R}$ is some number x_0 so that every $x \in S$ satisfies $x \leq x_0$. A *least upper bound*, also called a *supremum* for S is an upper bound, say x_1, so that no upper bound is less then x_1. It is a well known fact that every set of real numbers S which is bounded from above has a unique supremum, denoted

$$\sup_{x \in S} x.$$

Use this fact to prove that every Cauchy sequence in $\mathbb{R}$ converges.

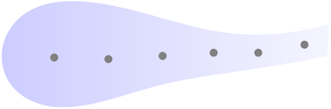
1.17 Use the result of problem 1.16 on the facing page to prove that every Cauchy sequence in $\mathbb{R}^n$ converges.

1.18 Use the result of problem 1.17 to prove that every Cauchy sequence in any closed subset of $\mathbb{R}^n$ converges to a point of that closed subset.

1.19 If $A \subset \mathbb{R}^n$ is any set, prove that there is a unique closed set, usually denoted $\bar{A}$, the *closure* of A, which is closed, contains A, and lies in any closed set containing A.

If $f \colon X \to Y$ is any map, the *support* of f is the closure of the set of points where $f \neq 0$. We say that f is supported on a set S to mean that S contains the support of f.

Compactness



Jumping farther and farther away in an unbounded set.

Jump around, but try to avoid stepping in, or near, the same place twice. If you live inside $\mathbb{R}$, you can just keep jumping forward along the axis; you never end up anywhere near where you were before. The same is true for any unbounded set. But if you live on a bounded set, you can only get so far away from where you were before; eventually you jump close to a point you have already jumped on. If you live in an open ball, you can at least try to head toward the boundary. That way you are always eventually headed away from any point of the ball, so the points you land on don't converge to a point of the ball: only toward a point of the boundary, which is outside the ball. Or you could choose a sequence of points that jump around different directions in the ball, getting closer to the boundary, but not converging to anything.

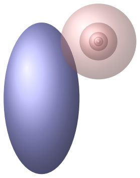
Jump around on the set $S = [0, 1]$. Your steps might not converge, because you just hop back and forth, say $0, 1, 0, 1, \dots$, or maybe $0, .9, 0, .99, \dots$ or $0, .5, 1, .5, 0, .5, 1, \dots$ or something similar to that. No matter how you choose to jump, write down the first decimal digit of each point you land at (say, after rounding off to the nearest digit). If you take infinitely many jumps, since there are only 10 possible decimal digits, infinitely many of your jumps must land on points with the same first decimal digit. In fact, infinitely many jumps must share the same first two decimal digits, and so on. By induction, we see that there must be an infinite collection of your jumps converging to some point in the set $S = [0, 1]$. On the other hand, if you jump in $S = (0, 1)$, you can avoid this fate: you can converge to 1, which doesn't lie in S.

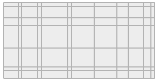
A set of points $S \subset \mathbb{R}^n$ is *compact* if every infinite sequence of points $x_1, x_2, \dots$ of S has a subsequence, like maybe $x_{12}, x_{3007}, \dots$, which converges.

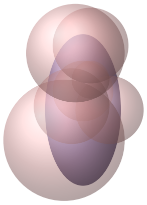
Theorem 1.1 (Heine–Borel)**.** *A subset $S \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. Suppose that S is compact. If S is not bounded, we can choose points of S, each many times farther away than the last, and jump away along them. Any subsequence will fly away even faster.

Suppose that S is not closed. So the complement $T = \mathbb{R}^n - S$ is not open. For T to be open means that each point of T lies in a ball inside T. So for T to fail to be open means that there is a point $x \in T$ so that every ball around x does not entirely lie in T. In other words, every ball around x contains a point not in T, i.e. a point in S. So if we pick balls $B_1, B_2, \dots$ around x, with radii going to zero, we can find points $x_1, x_2, \dots \in S$ with $x_1 \in B_1$, $x_2 \in B_2$, so converging to x. Any subsequence will converge even faster to x. But some subsequence must converge to a point of S, so $x \in S$, a contradiction. So S is closed.


Red balls surrounding a point outside a blue compact set

Suppose that S is closed and bounded, and let's prove that S is compact. Pick a sequence $x_1, x_2, \dots \in S$. We will let the reader prove that, since S is bounded, we can find a box containing S. Divide up the box into a grid of boxes, overlapping only on their faces, like this:  We let the reader fill in a precise definition of such a grid. We can make sure that all of the grid boxes have sides of length less than any amount we like. Infinitely many of the points of our sequence must lie in the same grid box, because there are only finitely many grid boxes. We repeat this argument, dividing that one grid box into a finer grid, with all sides much smaller than the previous grid, and again find infinitely many points in one of those grid boxes. As our subsequence, write down the first point in our sequence that lies in the first chosen grid box, and then write down the first point in our sequence that lies in the second, much smaller, chosen grid box (inside the first one), and so on. If we make each grid much smaller than the last, say with side lengths getting smaller by a factor of 10, then the first so many of the decimal digits of the points we pick agree, and the number of digits agreeing gets larger and larger. So the subsequence of these points converges. Because S is closed, the subsequence converges to a point of S. $\square$


A covering of a blue ellipsoid by a collection of red balls

An *open cover* of a set $S \subset \mathbb{R}^n$ is a collection of open sets U_α whose union contains every point of S:

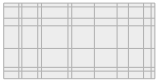
$$S \subset \bigcup_\alpha U_\alpha.$$

The plane is covered by the open balls. If you throw away some of the sets inan open cover, and still manage to have an open cover left over, then the left over sets form a *subcover*. For example, the plane is covered by all of the open balls of radius 2, but also covered by all of the open balls of radius 2 centered at the points with integer coordinates. We are just discarding certain of the sets from the open cover, *not* making the sets smaller.

1.20 What is the smallest number r so that all points of $\mathbb{R}^n$ lie a distance $\leq r$ from a point with integer coordinates?

Any open cover of a finite set of points has a finite subcover: we just need one open set covering each point.

Theorem 1.2 (Heine–Borel)**.** *A subset $S \subset \mathbb{R}^n$ is compact if and only if every open cover has a finite subcover.*

Proof. Suppose that S is closed and bounded. Take an open cover of S, say by open sets U_α. We will let the reader prove that, since S is bounded, we can find a box, say X_0, containing S. Divide up that box X_0 into a grid of boxes, overlapping only on their faces, like this:  We let the reader fill in a precise definition of such a grid. We can ensure that all of the sides of all of the grid boxes have length less than some number.

Call the various boxes of our grid $Y_1, Y_2, \dots, Y_N$. For each grid box Y_j, try to cover $S \cap Y_j$ by a finite number of the open sets U_α of the open cover. Suppose that you alway succeed. Then you put these finitely many subcovers together to form a finite subcover of S. So we can suppose that at least one of the sets $S \cap Y_j$ has no finite subcover by the open sets U_α. Let X_1 be a new name for Y_j, and pick a point $x_1 \in X_1$. So inside X_0 we have another box X_1, with sides all less than some number, and we have a point $x_1 \in X_1$.

Repeat infinitely often, generating boxes $X_1, X_2, X_3, \dots$. Arrange that each X_j has all of its sides of length less than some bound, and pick a point $x_j \in X_j$. As you do so, choose the bounds on lengths of sides so that they go to zero. In particular, the side lengths of X_j vanish as j gets large. Clearly the sequence $x_1, x_2, \dots$ is a Cauchy sequence, so converges, say to some point x. By the result of problem 1.18 on page 5, $x \in S$. Pick some open set U_{α_0} from the open cover containing x. Then U_{α_0} contains a ball around x, and therefore already contains every X_j, for large enough j. But then X_j is covered by finitely many of the U_α (even by just one of them), a contradiction. □

1.21* Finish the proof of the Heine–Borel theorem: prove that a set $S \subset \mathbb{R}^n$ with the property that every open cover has a finite subcover is closed and bounded.

1.22 Any open ball can be covered by finitely many open balls (even by just one ball: itself). Does that ensure that any open ball is compact? Explain.

1.23 Let I be the interval consisting of the real numbers x so that $0 < x < 1$. Exhibit an open cover of I with no finite subcover. So I is not compact.

Lemma 1.3. *Suppose that $f \colon K \to \mathbb{R}^q$ is a continuous map on a compact set $K \subset \mathbb{R}^p$. Then we can uniformly control how bad the error in approximating the output is from knowing how bad the error in approximating the input is; to be precise, for any $\varepsilon > 0$, there is some $\delta > 0$, so that for every $x \in K$ and $x + h \in K$, if $\|h\| < \delta$, then $\|f(x+h) - f(x)\| < \varepsilon$.*

The point is that we don't need to use a different δ for different points x.

Proof. For the rest of the proof, fix a value of $\varepsilon > 0$. By continuity, for each point x, we can choose some $\delta = \delta(x)$ to get this to work, i.e. so that for any h, if $\|h\| < \delta$, then $\|f(x+h) - f(x)\| < \varepsilon/2$.

For each point $x \in K$, let B_x be the ball of radius $\delta(x)/2$ about x; the sets B_p form an open cover of K. Since K is compact, there is a finite subcover, say by balls $B_{x_1}, B_{x_2}, \dots, B_{x_N}$. Let δ_0 be the minimum radius of these balls. (Crucial fact: $\delta_0 > 0$, since there are only finitely many balls.)

Take any points $x \in K$ and $x + h \in K$ so that $\|h\| < \delta_0$. Since these balls $B_{x_1}, B_{x_2}, \dots, B_{x_N}$ cover K, x lies in one of them, say in B_{x_j}. So $\|x - x_j\| < \delta\left(x_j\right)/2$. Therefore

$$\left\|f(x) - f\left(x_j\right)\right\| < \varepsilon/2.$$

But then since $\|h\| < \delta_0$, we also have $\|x + h - x\| = \|h\| < \delta_0$,

$$\left\|x + h - x_j\right\| < 2\delta_0 \le \delta\left(p_j\right),$$

so

$$\left\|f(x+h) - f\left(x_j\right)\right\| < \varepsilon/2.$$

By the triangle inequality,

$$\begin{aligned} \|f(x+h) - f(x)\| &\le \left\|f(x+h) - f\left(x_j\right)\right\| + \left\|f\left(x_j\right) - f(x)\right\|, \\ &< \varepsilon. \end{aligned}$$

□

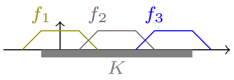
1.24* Prove that every continuous function on a compact set takes on a minimum value and a maximum value. Does it have to take on all values in between?

Partitions of unity

1.25 Give an example of a continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ so that $f(x) = 1$ if $\|x\| \le 1$ and $f(x) = 0$ if $\|x\| \ge 2$.

A *partition of unity* on a compact set $K \subset \mathbb{R}^n$ is a collection of continuous functions $f_1, f_2, \dots, f_N$ so that

1. $0 \le f_1, f_2, \dots, f_N \le 1$ and
2. $f_1(x) + f_2(x) + \dots + f_N(x) = 1$ for each $x \in K$.



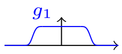
A partition of unity on a compact set K

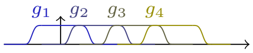
A partition of unity $f_1, f_2, \dots, f_N$ is *subordinate* to an open cover $\{U_\alpha\}$ if for every function f_i, there is some open set U_α from the open cover, so that f_i is supported on U_α.

Theorem 1.4. *For any open cover of any compact set, there is a partition of unity subordinate to it.*

Proof. Take $K \subset \mathbb{R}^n$ a compact set and an open cover of K by open sets U_α. We need to prove that there is a partition of unity on K subordinate to that open cover.

Each point $x \in K$ lies in one of these U_α, say $x \in U_{\alpha(x)}$. Since $U_{\alpha(x)}$ is open, we can pick an open ball B_x around x lying inside $U_{\alpha(x)}$. Shrinking B_x if needed, we can assume that $\bar{B}_x \subset U_{\alpha(x)}$. Shrinking even more, let $C_x \subset B_x$ be a smaller open ball. So $\bar{C}_x \subset B_x$ too. The various C_x, for $x \in X$, form an open cover. Pick a finite subcover, say $C_{x_1}, C_{x_2}, \dots, C_{x_N}$. For each $i = 1, 2, \dots, N$, pick a continuous function g_i so that $g_i = 1$ inside C_{x_i} and $g_i = 0$ outside B_{x_i}.





These g_i might not add up to 1. Let $f_1 = g_1$ and inductively let

$$f_{i+1} = (1 - g_1)(1 - g_2) \dots (1 - g_i)\, g_{i+1}.$$

By induction, check that

$$\begin{aligned} f_1 &= 1 - (1 - g_1), \\ f_1 + f_2 &= 1 - (1 - g_1)(1 - g_2), \\ &\vdots \\ f_1 + f_2 + \dots + f_N &= 1 - \prod_i (1 - g_i). \end{aligned}$$

On C_{x_j}, $g_j = 1$ so $f_1 + f_2 + \dots + f_N = 1$. The various C_{x_j} cover K, so $f_1 + f_2 + \dots + f_N = 1$ everywhere on K. ☐

Review problems

1.26 Prove that the *multiplication map*

$$f\colon \mathbb{R}^2 \to \mathbb{R}, f(x, y) = xy$$

is continuous.

1.27 Prove that every linear map is continuous.

1.28 Suppose that $f\colon \mathbb{R}^3 \to \mathbb{R}^2$. We can write f as

$$f(x, y, z) = \begin{pmatrix} f_1(x, y, z) \\ f_2(x, y, z) \end{pmatrix}.$$

Prove that f is continuous just when both of f_1 and f_2 are continuous.

1.29 Suppose that $A \subset \mathbb{R}^n$ is compact and $x \in \mathbb{R}^n$ and $x \notin A$. Prove that there is a positive lower bound to the distances between x and any of the points of A. Prove that this positive distance is acheived at some point of A.

1.30 Prove that if $f\colon \mathbb{R}^q \to \mathbb{R}^p$ is continuous and $A \subset \mathbb{R}^q$ is compact then $f(A)$ is also compact.

Notes

1. If I attack you with an ε, you can reach into your bag of δ's, and with one of them you can defend yourself.
2. The idea of compactness is subtle. Compactness is similar to finiteness. Imagine that at every point in some set, you station a Garda. But each Garda is nearsighted, and some are very nearsighted. You only need finitely many of them to police a compact set, as if the set were finite. But on a noncompact set, like an infinite line, you might need infinitely many. Even on a bounded set, like an open interval, if each Garda has less and less range of vision as you get closer to the boundary, you might need infinitely many of them.
3. The definition of compactness we used above is properly called *sequential compactness*. In more elaborate spaces than $\mathbb{R}^n$, it is convenient to define compactness by saying that a compact set is one on which every open cover has a finite subcover.
4. To see what happens when we don't have compactness, look at the function $f(x) = 1/x$ for $x > 0$. Continuity of $f(x)$ for $x > 0$ is precisely that fact that, if you fix some $\varepsilon > 0$ and a point $x > 0$, I can pull out a $\delta > 0$ somehow from my bag of δ's so that, for all h with $|h| < \delta$, I can get

   $$\left| \frac{1}{x+h} - \frac{1}{x} \right| < \varepsilon.$$

   We let the reader check that the biggest δ we can take is

   $$\delta = \frac{x}{1 + \frac{1}{x\varepsilon}}.$$

   The smaller the x we are faced with, the smaller the δ we have to employ, and we can't use the same δ for all $x > 0$. Of course, if we insist that x stays in a closed interval (which is compact!) where f is defined, say for example $1 \le x \le 2$, there is a smallest ("worst") δ we might need to use, and we can use that same δ for all of these x:

   $$\delta = \frac{1}{1 + \frac{1}{\varepsilon}}.$$
5. The concept of partition of unity we have defined is special to compact sets. The more general concept is a little more complicated, but we won't need it; see Spivak [7] p. 63.

Chapter 2

Differentiation

We learn how to differentiate maps of Euclidean spaces.

Definition

Let's write $f(x) = o(x)$ to mean $\frac{f(x)}{\|x\|} \to 0$ as $x \to 0$. In one variable calculus,

$$f(x+h) = f(x) + f'(x)h + o(h)$$

Lets try to use exactly the same equation in many variables. What could it mean?

Suppose that $f \colon \text{open } \subset \mathbb{R}^q \to \mathbb{R}^p$. In one variable calculus, $f'(x)$ is a number. But that doesn't make sense here. We have $x, h \in \mathbb{R}^q$ and $f(x), f(x+h) \in \mathbb{R}^p$. So $f'(x)$ must be something that eats a vector $h \in \mathbb{R}^q$ and spits out a vector $f'(x)h \in \mathbb{R}^p$. We guess then that $f'(x)$, for any fixed value of x, should be not a number but a linear map, i.e. a matrix.

So we say that a $p \times q$ matrix A is a *derivative* of a map $f \colon \text{open } \subset \mathbb{R}^q \to \mathbb{R}^p$ at a point x if

$$f(x+h) = f(x) + Ah + o(h) .$$

If f has a derivative A at a point x, then we say f is *differentiable* at x. If f has a derivative at every point where f is defined, we say that f is *differentiable.*

2.1 Prove that if A and B are matrices and both A and B are derivatives of a map $f(x)$ at at point x, then $A = B$. So from now on, we can refer to *the* derivative, instead of *a* derivative.

So from now on, we call the derivative $f'(x)$.

If $q = 1$, so f is a vector-valued function of a real variable, we can think the point $f(t)$ as travelling on a curve in space when we vary t, and t as a variable representing time. Then $f'(t)$ is a one-column matrix, which we always think of as a vector, not a matrix, the *velocity* of the curve.

2.2 Use the definition above to differentiate the map

$$(y_1, y_2) = f(x_1, x_2) = \left(x_1 + x_2, x_1 + x_2^2\right) .$$

2.3 Suppose that $f \colon \mathbb{R}^q \to \mathbb{R}^p$ is a linear map. Prove that f is differentiable and that $f'(x) = f$ for every point $x \in \mathbb{R}^q$.

2.4 Differentiate the function

$$(x, y) = f(r, \theta) = (r \cos\theta, r \sin\theta)\,.$$

2.5 Prove that if a map defined in an open set is differentiable at a point then it is continuous at that point.

2.6* Let f be the function associating to each $n \times n$ matrix M its inverse: $f(M) = M^{-1}$. Prove that f is defined on an open set. Find $f'(M)$ for any invertible matrix M. Hint: $f'(M)$ is a linear map taking matrices to matrices. Hint: use the identity

$$(I + h)^{-1} = I - h + h^2 + \dots$$

for any matrix h close enough to 0.

Partial derivatives

The *partial derivative* of a function $f(x, y)$ with respect to x is

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

(The notation $\frac{\partial f}{\partial x}$ is used instead of $\frac{df}{dx}$ to remind us that x is not the only variable.) Similar notation applies to more variables.

Lemma 2.1. *If f is a differentiable map, then*

$$f'(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_q} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_q} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \frac{\partial f_p}{\partial x_2} & \dots & \frac{\partial f_p}{\partial x_q} \end{pmatrix}.$$

Proof. Recall that if A is a $p \times q$ matrix, and $e_1, e_2, \dots, e_q$ are the standard basis vectors of $\mathbb{R}^q$, then Ae_1 is the 1st column of A, Ae_2 is the 2nd column of A, etc. Applying $f'(x)$ to the standard basis we find that for small real numbers h,

$$f'(x)he_j = f\left(x + he_j\right) - f(x) + o\left(h\right),$$

so that, dividing by h,

$$\begin{aligned} f'(x)e_j &= \lim_{h \to 0} \frac{f\left(x + he_1\right) - f(x)}{h} \\ &= \frac{\partial f}{\partial x_j}. \end{aligned}$$

□

Chain rule

Theorem 2.2. *Suppose that* $U \subset \mathbb{R}^r$, $V \subset \mathbb{R}^q$ *are open sets and* $f\colon U \to V$ *and* $g\colon V \to \mathbb{R}^p$ *are differentiable maps. Let* $F = g \circ f$, *i.e.* $F(x) = g(f(x))$ *for all* $x \in U$. *Then* $F'(x) = g'(f(x))f'(x)$, *where the right hand side is a matrix multiplication.*

Proof. As $h \to 0$,

$$\begin{aligned} F(x+h) - F(x) &= g(f(x+h)) - g(f(x)) \\ &= g\left(f(x) + f'(x)h + o\left(h\right)\right) - g(f(x)) \\ &= g\left(f(x)\right) + g'(f(x))f'(x)h + o\left(h\right) - g(f(x)) \\ &= g'(f(x))f'(x)h + o\left(h\right). \end{aligned}$$

□

Uniform estimates

2.7 Suppose that $f(x, y)$ is a function defined on an open subset of the plane, continuous in x for any fixed y, and continuous in y for any fixed x. Prove that $f(x, y)$ is continuous.

Suppose that $U \subset \mathbb{R}^q$ is an open set. A map $f\colon U \to \mathbb{R}^p$ is *continuously differentiable* if the map $f'\colon U \to \mathbb{R}^{p\times q}$ is continuous. A map defined on any set in $\mathbb{R}^n$ (perhaps not open) is *continuously differentiable* if near each point it can be extended to a continuously differentiable function on some open set.

Theorem 2.3. *A map* $f\colon$ *open* $\subset \mathbb{R}^p \to \mathbb{R}^q$ *is continuously differentiable just when all of the partial derivatives* $\frac{\partial f_i}{\partial x_j}$ *are defined and continuous.*

Proof. It is enough to check for each component function $f_1, f_2, \dots, f_q$, so we can assume that $q = 1$, i.e. $f\colon$ open $\subset \mathbb{R}^p \to \mathbb{R}$ is a real-valued function.

By lemma 2.1 on the preceding page, if f is continuously differentiable, then the partial derivatives are defined and continuous. Suppose that the partial derivatives $\frac{\partial f}{\partial x_j}$ are defined and continuous, and we will prove that f is continuously differentiable. Pick a vector h, with entries small enough that the box with corners x and $x + h$ lies inside the open set where f is defined. Write h as

$$h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}.$$

Let

$$k_0 = 0, k_1 = \begin{pmatrix} h_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, k_2 = \begin{pmatrix} h_1 \\ h_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, k_n = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \\ \vdots \\ h_n \end{pmatrix} = h.$$

Then

$$f\left(x + k_{j+1}\right) - f\left(x + k_j\right) = \frac{\partial f}{\partial x_{j+1}}\left(x + k_j\right) h_{j+1} + o\left(h_{j+1}\right).$$

So

$$\begin{aligned} f(x+h) - f(x) &= \sum_{j=0}^{n} f\left(x + k_{j+1}\right) - f\left(x + k_j\right), \\ &= \sum_{j=0}^{n} \frac{\partial f}{\partial x_{j+1}}\left(x + k_j\right) h_{j+1} + o(h), \\ &= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\left(x + k_{j-1}\right) h_j + o(h). \end{aligned}$$

If we make all entries of h small enough, then by continuity we can approximate $\frac{\partial f}{\partial x_j}\left(x + k_{j-1}\right)$ as closely as we like by $\frac{\partial f}{\partial x_j}(x)$, say within some value ε. So for any $\varepsilon > 0$, we can arrange that

$$\left| f(x+h) - f(x) - \sum_j \frac{\partial f}{\partial x_j} h_j \right| < \varepsilon \left\| h \right\|,$$

i.e.

$$f(x+h) - f(x) - \sum_j \frac{\partial f}{\partial x_j} h_j = o(h).$$

□

Lemma 2.4. *Suppose that* $U \subset \mathbb{R}^q$ *is an open set and* $f \colon U \to \mathbb{R}^p$ *is a continuously differentiable map and* $K \subset U$ *is a compact set. Then we can control the remainder term in the derivative uniformly over all of* K*. To be precise, let*

$$Q(x,h) = f(x+h) - f(x) - f'(x)h,$$

the "remainder". Then for any $\varepsilon > 0$*, there is a* $\delta > 0$ *so that, for any* $x \in K$*, and any* h *with* $\|h\| < \delta$*,*

$$\|Q(x,h)\| < \varepsilon \left\| h \right\|.$$

(Instead of having to work with a different δ *at each point* x*, you get to use the same* δ *at every* x *throughout* K*.)*

Proof. The function

$$R(x,h) = \begin{cases} \frac{Q(x,h)}{\|h\|}, & \text{if } h \neq 0, \\ 0, & \text{if } h = 0 \end{cases}$$

is clearly continuous; apply lemma 1.3 on page 7. ☐

Bump functions

2.8 Let

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x > 0, \\ 0, & \text{if } x \leq 0. \end{cases}$$

Prove that f is continuously differentiable and that $0 = f(0) = f'(0)$. Hint: try l'Hôpital's rule.

2.9 Let

$$f(x) = \begin{cases} e^{-1/x^2} e^{-1/(x-1)^2}, & \text{if } 0 < x < 1, \\ 0, & \text{if } x \geq 1 \text{ or } x \leq 0. \end{cases}$$

Prove that f is continuously differentiable and both f and f' vanish at $x = 0$ and at $x = 1$.

2.10 Prove that there is a continuously differentiable function $g \colon \mathbb{R} \to \mathbb{R}$ so that $g(x) = 0$ for $x \leq 0$ and $g(x) = 1$ for $x \geq 1$ and g is strictly increasing.

2.11 Take a function g as in problem 2.10, and pick any numbers $R > r > 0$. Let

$$k(x) = g\left(\frac{R - x}{R - r}\right).$$

Draw pictures. Why is $k(x) = 1$ for $0 \leq x \leq r$ and $k(x) = 0$ for $x \geq R$? Why is $k(x)$ decreasing and continuously differentiable?

Following the previous problem, define a function $h \colon \mathbb{R}^n \to \mathbb{R}$ by $h(x) = k\left(\|x - x_0\|\right)$, for some choice of point x_0 and radii $r < R$. Clearly $h = 1$ inside the ball of radius r around x_0, and $h = 0$ outside the ball of radius R. Moreover, h is continuously differentiable, and strictly decreasing as we increase $\|x\|$. We will call such a function h a *bump function*.

2.12 Suppose that $x \colon [0,1] \to \mathbb{R}^n$ and $y \colon [0,1] \to \mathbb{R}^n$ are continuously differentiable curves. Suppose that $x(1) = y(0)$. Prove that there is a continuously differentiable curve $z \colon [0,1] \to \mathbb{R}^n$ which "travels along x" and then "travels along y". Hint: imagine that x and y meet at an angle. Hint: slow down as you get to the end of x, and then start very slowly on y.

Higher derivatives and optimization

This section can be omitted without loss of continuity.

Write

$$\frac{\partial^2 f}{\partial x \partial y} \text{ to mean } \frac{\partial}{\partial x}\frac{\partial f}{\partial y},$$

and so on. The *second derivative* of a function $f \colon \text{open } \subset \mathbb{R}^n \to \mathbb{R}$ is the matrix

$$f''(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_1} \\ \frac{\partial^2 f}{\partial x_2 \partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

Write $o\left(h\right)^2$ to mean a quantity $Q(h)$ so that

$$\frac{Q(h)}{\|h\|^2} \to 0 \text{ as } h \to 0.$$

2.13 To acclimatise to the notation, expand out into partial derivatives to show that if f' is differentiable, then $f'(x) - f'(0) = x^t f''(0) + o\left(x\right)$, where x^t is the transpose of the vector x. Why isn't it $f''(0)x + o\left(x\right)$?

Lemma 2.5. *If $f \colon$ open $\subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point x_0 and f' is also differentiable at x_0, then $f''\left(x_0\right)$ is a symmetric matrix.*

Proof. For simplicity of notation, assume that $x_0 = 0$. Pick two vectors $h, k \in \mathbb{R}^n$ so that the parallelogram with corners at $0, h, k, h+k$ lies inside the domain where f is defined. Let

$$g(t) = f\left(th + k\right) - f\left(th\right).$$

Expand out:

$$g'(t) = \left(f'\left(th + k\right) - f'\left(th\right)\right) h,$$

and if we write v^t for the transpose of any vector v, we can write this as

$$\begin{aligned} &= \Big(f'\left(th + k\right) - f'\left(0\right) - \left(th + k\right)^t f''(0)\Big) h \\ &\quad - \left(f'\left(th\right) - f'\left(0\right) - th^t f''(0)\right) h \\ &\quad - k^t f'(0). \end{aligned}$$

By differentiability of f', the first two terms are $o\left(h, k\right) o\left(h\right)$.

By the remainder formula for the Taylor series in one variable calculus (which we won't prove),

$$\left|g(1) - g(0) - g'(0)\right| \le \max_{0 \le t \le 1} \left|g'(t) - g'(0)\right|.$$

Therefore

$$g(1) - g(0) - k^t f''(0)h = o\left(h,k\right)^2 .$$

But $g(1) - g(0)$ is symmetric in h and k, so swapping roles of h and k, we get the same result, and therefore

$$h^t f''(0)k - k^t f''(0)h = o\left(h,k\right)^2 .$$

□

Theorem 2.6. *Suppose that f is a twice continuously differentiable function, i.e. f, f' and f'' are defined and continuous, on some open set in $\mathbb{R}^n$. If f is defined on the line segment between two points x and $x+h$, then for some number t with $0 \le t \le 1$,*

$$f\left(x+h\right) = f\left(x\right) + f'\left(x\right)h + \frac{1}{2}h^t f''\left(x+th\right)h.$$

Proof. If we let $g(s) = f(x+sh)$, then we apply the chain rule to compute the Taylor series for $g(s)$ and use the Taylor remainder theorem from one variable calculus. □

Theorem 2.7. *Suppose that f is a twice continuously differentiable function, i.e. f, f' and f'' are defined and continuous, on some open set in $\mathbb{R}^n$. If f is defined on an open ball about a point x, and $x+h$ lies in that ball, then*

$$f\left(x+h\right) = f\left(x\right) + f'\left(x\right)h + h^t f''\left(x\right)h + o\left(h\right)^2 .$$

Proof. By theorem 2.6,

$$f\left(x+h\right) = f\left(x\right) + f'\left(x\right)h + \frac{1}{2}h^t f''\left(x\right)h + Q(x,h),$$

where

$$Q(x,h) = \frac{1}{2}h^t f''\left(x+th\right)h - \frac{1}{2}h^t f''\left(x\right)h,$$

for some $t = t\left(x,h\right)$ with $0 \le t \le 1$. By continuity of f'', $Q(x,h) = o\left(h\right)^2$. □

2.14 Suppose that A is a matrix, and $g(x) = f(Ax)$ for all x. Prove that $g'(x) = f'(Ax)A$ and that $g''(x) = A^t f''(Ax)A$.

Theorem 2.8. *Suppose that $f\colon$ open $\subset \mathbb{R}^n \to \mathbb{R}$ is a function whose first and second derivatives, i.e. $\frac{\partial f}{\partial x_i}$ and $\frac{\partial f}{\partial x_i \partial x_j}$, are defined and continuous. If f has a minimum at a point x_0, then $f'\left(x_0\right) = 0$ and each eigenvalue λ of the matrix $f''\left(x_0\right)$ has $\lambda \ge 0$. On the other hand, if x_0 is a point at which $f'\left(x_0\right) = 0$ and each eigenvalue λ of $f''\left(x_0\right)$ is positive, $\lambda > 0$, then x_0 is a local maximum for f, and in fact a strict local maximum: $f(x) < f\left(x_0\right)$ for all x sufficiently close to, but not equal to, x_0. The same reasoning works, with obvious modifications, for finding local maxima.*

Proof. We leave it to the reader to prove that $f'(x_0) = 0$ at a minimum. As is well known in linear algebra, a symmetric matrix can be diagonalized by rotation of the choice of coordinate axes. Rotate the choice of coordinate axes around the point x_0 to arrange that $f''(x_0)$ is diagonal. From one variable calculus, working one variable at a time, we see that if f has a maximum at x_0, then $f''(x_0)$ can not have negative eigenvalues.

Suppose that $f'(x_0) = 0$ and that the eigenvalues of $f''(x_0)$ are positive. For simplicity, we can assume that $f''(x_0)$ is a diagonal matrix with only positive diagonal entries. We want to prove that x_0 is a strict local minimum. For simplicity of notation, assume that $x_0 = 0$ and that $f(0) = 0$. We need to prove that $f(x) > 0$ near $x = 0$. Writing

$$\begin{aligned} f(x) &= f(0) + f'(0)x + x^t f''(0)x + o(x)^2, \\ &= x^t f''(0)x + o(x)^2. \end{aligned}$$

Expanding out, if

$$f''(0) = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

then

$$x^t f''(0)x = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \dots + \lambda_n x_n^2.$$

In particular, $x^t f''(0)x \ge \lambda \|x\|^2$, where $\lambda > 0$ is the smallest eigenvalue of $f''(0)$. So

$$\frac{f(x)}{\|x\|^2} \ge \lambda + \frac{o(x)^2}{\|x\|^2} > 0$$

for all $x \neq 0$ close enough to 0. ☐

For example, consider the function

$$f(x, y) = \frac{1}{x^2 y^2} + x^2 + y^2,$$

defined on the plane, except on the two axes $x = 0$ and $y = 0$. Calculate

$$\frac{\partial f}{\partial x} = 2x - \frac{2}{x^3 y^2}, \quad \frac{\partial f}{\partial y} = 2y - \frac{2}{x^2 y^3}.$$

Clearly these vanish just when

$$x = \frac{1}{x^3 y^2} \text{ and } y = \frac{1}{x^2 y^3},$$

or in other words, clearing denominators,

$$x^4 = y^2 \text{ and } y^4 = x^2.$$

Therefore, $y^2 = x^4 = y^8$, and so $y = \pm 1$ and $x = \pm 1$. So there are 4 points where $f'(x, y) = 0$ (*critical points*): $(1, 1), (1, -1), (-1, 1), (-1, -1)$. Calculate that

$$f''(x, y) = \begin{pmatrix} 2 + \frac{6}{x^4 y^2} & \frac{4}{x^3 y^3} \\ \frac{4}{x^3 y^3} & 2 + \frac{6}{x^2 y^4} \end{pmatrix},$$

so that at $(x, y) = (\pm 1, \pm 1)$, we find

$$f''(x, y) = \begin{cases} \begin{pmatrix} 8 & -4 \\ -4 & 8 \end{pmatrix}, & \text{if } x = y\ , \\ \begin{pmatrix} 8 & 4 \\ 4 & 8 \end{pmatrix}, & \text{if } x \neq y\ . \end{cases}$$

In either case, check that the eigenvalues of $f''(x, y)$ at $(x, y) = (\pm 1, \pm 1)$ are $\lambda = 4$ and $\lambda = 12$. So $f''(x, y)$ at $(x, y) = (\pm 1, \pm 1)$ has all eigenvalues positive, so each of the 4 points $(x, y) = (\pm 1, \pm 1)$ is a local minimum of f.

2.15 Find all local maxima and local minima of

$$f(x, y) = 3x^2 + 2x^3 + 2y^3 - 3y^2.$$

2.16 Explain why the (local or global) maxima and minima of the function

$$f(x, y) = e^{x^2+y^2+3xy}$$

occur precisely where the corresponding maxima and minima of

$$g(x, y) = x^2 + y^2 + 3xy$$

occur. Find these local maxima and minima, and determine which are merely local maxima and minima and which are truly global maxima or minima.

2.17 Prove that the function

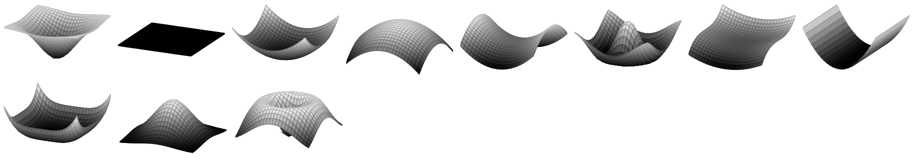
$$f(x, y) = \left(y - 3x^2\right)\left(y - x^2\right)$$

has a local minimum on every straight line through the origin (i.e. if we drag the point (x, y) along a curve of the form $(x, y) = t(x_0, y_0)$), but that $f(x, y)$ does not have a local minimum at the origin.

2.18 Draw $f(x, y) = x^2 - y^2$, and compute $f''(0)$.

2.19 A *saddle point* is a point where f'' has eigenvalues of different signs. Prove that arbitrarily close to a saddle point of a function, there are points where the function takes on a larger value and points where it takes on a smaller value.

2.20 Which plot belongs with which function? Explain how you can see the signs of the eigenvalues of the second derivative in the plot. From the plots, which pairs of functions have the same pattern of + and − signs and zeroes occuring in the eigenvalues of their second derivatives?



$$\left(x^2+y^2\right)e^{-x^2-y^2},\ x^2,\ x^2+y^3,\ e^{-x^2-y^2},\ -\left(x^2+y^2\right)e^{-x^2-y^2},\ x^2-y^2,$$
$$-x^2-y^2,\ x^4+y^4,\ x^2+y^2,\ 1,\ -e^{x^2+y^2}$$

Review problems

2.21 Suppose that α is a real number. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be the function $f(x,y) = |xy|^{\alpha}$. For which values of α is f differentiable at $(x,y) = (0,0)$?

2.22 Suppose that $f\colon \text{open} \subset \mathbb{R}^n \to \mathbb{R}$ is a differentiable function and that $f(x)$ has a minimum at $x = x_0$. Prove that $f'(x_0) = 0$.

2.23 Suppose that $x\colon \mathbb{R} \to \mathbb{R}^3$ is a continuously differentiable map and that $\|x'(t)\| = 1$ for all t. Prove that $\langle x'(t), x(t)\rangle = 1$ for all t. Draw a picture to explain why.

2.24 Give an example of a continuously differentiable function $f(x,y)$ defined on a connected open set in the plane so that $\frac{\partial f}{\partial y} = 0$ but f is not expressible as a function of x. Hint: take an open set shaped like a horseshoe.

2.25 Prove that every differentiable map is continuous.

2.26* Find $\det'$: let $f\colon \mathbb{R}^{n^2} \to \mathbb{R}$ be the map $f(M) = \det(M)$. Prove that if M is invertible, then

$$f'(M)N = \det(M)\operatorname{tr}\left(M^{-1}N\right).$$

Notes

1. We call $f'(x)$ the *derivative* of $f(x)$, but it is often called the *Jacobian* of $f(x)$, especially in older mathematical literature. The term *Jacobian* is, confusingly, also used to mean the determinant $\det f'(x)$.
2. If $f\colon \text{open} \subset \mathbb{R}^q \to \mathbb{R}^p$ has continuous partial derivatives $\frac{\partial f}{\partial x_i}$ in all of the variables $x_1, x_2, \dots, x_q$, then f is differentiable. On the other hand, there are examples of maps f for which those partial derivatives exist but are not continuous and which are not differentiable. Such counterexamples are not of much interest for us.

Chapter 3

The Contraction Mapping Principle

We prove that every contracting map has a unique fixed point. We will later use this to prove the inverse function theorem.

Contraction mappings in one variable

What happens when you keep squishing something by the same squishing map repeatedly? Maybe it squishes down to a point. Pick a number c so that $0 \le c < 1$. Suppose that $y = f(x)$ is a real-valued function of a real variable. Say that f is a *contraction map* with dilation at worst c if, for any two points $s, t \in \mathbb{R}$ $|f(s) - f(t)| \le c\,|s - t|$.

3.1 Prove that every contraction map is continuous.

3.2 Which of the following are contraction maps?
(a) $f(x) = x$
(b) $f(x) = 2x$
(c) $f(x) = x/3$
(d) $f(x) = e^x$
(e) $f(x) = \sin(x)$

3.3 Suppose that c is a number with $0 \le c < 1$. Suppose that $f \colon \mathbb{R} \to \mathbb{R}$ is differentiable and $|f'(x)| < c$. Prove that f is a contraction map with dilation c.

3.4 Suppose that c is a number with $0 \le c < 1$. Suppose that $f \colon \mathbb{R} \to \mathbb{R}$ is a contraction map with dilation c and that $f(x)$ is differentiable. Prove that $|f'(x)| < c$ for every real number x.

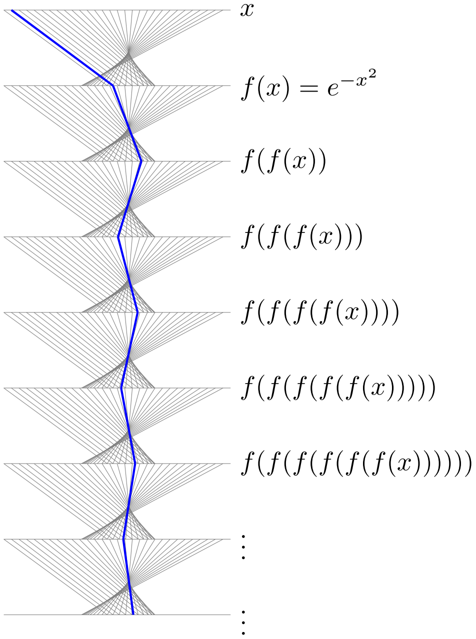
A *fixed point* of a function f is a number x_0 so that $f(x_0) = x_0$.

3.5 Find all fixed points of
(a) $f(x) = x$
(b) $f(x) = 2x$
(c) $f(x) = x/3$
(d) $f(x) = e^x$
(e) $f(x) = \sin(x)$

3.6 Prove that $f(x) = e^{-x^2}$ has a fixed point. (I don't know what it is.)

3.7 Let $f(x) = e^{-x^2}$. Draw the graph of $y = x$, $y = f(x)$, $y = f(f((x))$ and $y = f(f(f(x))$. Draw where you think the fixed point is.

3.8 How many fixed points can a polynomial of degree 100 have? How few? Prove your answer.



Theorem 3.1. *Suppose that* $f : \mathbb{R} \to \mathbb{R}$ *is a contraction map. Then* f *has a unique fixed point. If* x_0 *is any point, and we let* $D = |x_0 - f(x_0)|$, *then the fixed point is at most* $D/(1-c)$ *away from* x_0.

Proof. Just pick any point $x_0 \in \mathbb{R}$, and let

$$\begin{aligned} x_1 &= f(x_0), \\ x_2 &= f(x_1), \\ &\vdots \end{aligned}$$

etc. We want to prove that this sequence has a fixed point as limit. Imagine for the moment that we know the sequence is Cauchy. Then it must converge, say to some point x, so

$$x = \lim_{i \to \infty} f(x_i).$$

The function f is continuous, so

$$f\left(\lim_{j \to \infty} x_j\right) = \lim_{j \to \infty} f(x_j).$$

Therefore $f(x) = x$, a fixed point.

So we only need to prove that $x_0, x_1, x_2, \ldots$ is a Cauchy sequence. Let D be the distance from x_0 to x_1. How fast do the distances shrink? For example,

$$\begin{aligned} |x_1 - x_2| &= |f(x_0) - f(x_1)| \\ &\le c\,|x_0 - x_1| \\ &= cD. \end{aligned}$$

By induction, succesive points of the sequence are at distance

$$|x_i - x_{i+1}| \le c^i\,|x_0 - x_1|\,.$$

What about points further down the sequence?

$$\begin{aligned} |x_0 - x_2| &\le |x_0 - x_1| + |x_1 - x_2| \\ &= D + cD. \end{aligned}$$

Similarly by induction, if $i < j$, then

$$|x_i - x_j| \le c^i \left(D + cD + \cdots + c^{j-i-1}D\right).$$

Summing a geometric series, this is at worst

$$\begin{aligned} |x_i - x_j| &\le c^i\,(D + cD + \cdots +) \\ &\le c^{\min(i,j)} D \frac{1}{1-c}. \end{aligned}$$

Note that c and D are fixed. If we make both i and j large enough, $c^{\min(i,j)} \to 0$, so the sequence is Cauchy. Moreover, look what happens if we set $i = 0$ and $j \to \infty$: if x is the fixed point, then its distance from x_0 is at most the sum of that geometric series, $D/(1-c)$. □

Continuity

Theorem 3.2. *Suppose that $f(s,x)$ is a continuous function of two variables. Write $f_s(x)$ to mean $f(s,x)$. Suppose that for each s, the map f_s is a contraction map, with dilation no worse than some value $c(s)$ depending continuously on s. For each fixed value of s, let $x(s)$ be the fixed point of the map f_s. Then $x(s)$ is a continuous function as we vary s.*

Proof. Suppose that we take some value s_0 and some nearby value s_1. Let $x_0 = x(s_0)$ and $x_1 = x(s_1)$.

By continuity of $c(s)$, we can find a constant C with $0 \le c(s) < C < 1$ for all s close enough to s_0. So we can assume that for every s_1 with $|s_1 - s_0| < \delta$, we have $c(s_1) < C$.

Pick any $\varepsilon > 0$. We need to prove that by making s_1 close enough to s_0, we can ensure that $|x_1 - x_0| < \varepsilon$. Let $\varepsilon_1 = (1-C)\varepsilon$. Picking a small enough δ, and ensuring that s_1 is chosen close enough to s_0, say $|s_1 - s_0| < \delta$, will ensure (by continuity of $f(x, s)$ in the s variable) that

$$|f(s_0, x_0) - f(s_1, x_0)| < \varepsilon_1.$$

From theorem 3.1 on page 22, we see that the fixed point x_1 of f_{s_1} is within distance $D/(1 - c(s_1))$ of x_0. This number D is

$$\begin{aligned} D &= |x_0 - f(s_1, x_0)| \\ &\leq |x_0 - f(s_0, x_0)| + |f(s_0, x_0) - f(s_1, x_0)| \\ &= 0 + |f(s_0, x_0) - f(s_1, x_0)| \\ &\leq \varepsilon_1. \end{aligned}$$

Clearly $D/(1 - c(s_1)) < D/(1 - C)$, so

$$|x_1 - x_0| < \frac{\varepsilon_1}{1 - C} = \varepsilon.$$

□

How to test for contraction

Theorem 3.3. *Suppose that $y = f(x)$ is a differentiable real-valued function of a real variable, continuously differentiable except at a discrete set of points. Suppose that c is a constant real number so that $0 \leq c < 1$ and that $|f'(x)| < c$ for all x. Then f is a contraction map with dilation at most c.*

Proof. By the fundamental theorem of calculus, for any real numbers s and t,

$$f(t) - f(s) = \int_0^1 f'(ut + (1 - u)s)(t - s)\, du,$$

so

$$\begin{aligned} |f(t) - f(s)| &= \left| \int_0^1 f'(ut + (1 - u)s)(t - s)\, du \right| \\ &\leq \int_0^1 |f'(ut + (1 - u)s)|\, |t - s|\, du \\ &\leq \int_0^1 c\, |t - s|\, du \\ &= c\, |t - s| \end{aligned}$$

and similarly if $s \geq t$. □

Contraction mappings in Euclidean spaces

At this stage, we can restate the notations and theorems of this chapter in any number of variables (i.e. any number of dimensions). The proofs of the theorems are identical, just changing notation from absolute value symbols $|x|$ to Euclidean norms $\|x\|$.

Suppose that $X \subset \mathbb{R}^q$ and $Y \subset \mathbb{R}^p$ are two sets. Pick a number c so that $0 \le c < 1$. A map $f\colon X \to Y$ is a *contraction map* with *dilation* at worst c if, for any two points $s, t \in \mathbb{R}$

$$\|f(s) - f(t)\| \le c \|s - t\|.$$

If $X \subset \mathbb{R}^p$ is a set of points and $f\colon X \to X$ is a map, then a *fixed point* of f is a point $x_0 \in X$ so that $f(x_0) = x_0$.

Theorem 3.4 (Contraction mapping theorem)**.** *If $X \subset \mathbb{R}^p$ is a closed set of points, and $f\colon X \to X$ is a contraction map, then f has a unique fixed point.*

Theorem 3.5 (Continuous dependence)**.** *Suppose that $X \subset \mathbb{R}^p$ is a closed set of points, that $S \subset \mathbb{R}^q$ is an arbitrary set of points, and $f\colon S \times X \to X$ is a continuous map. Let $f_s(x) = f(s,x)$. Suppose that, for every $s \in S$, f_s is a contraction map with dilation at worst $c(s)$, where $0 \le c(s) < 1$ and $c\colon S \to \mathbb{R}$ is continuous. Then each map f_s has a unique fixed point, say $x(s)$, given by a continuous map $x\colon S \to X$.*

The *norm* of a $p \times q$ matrix A is

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

Warning: norms of matrices are hard to calculate.

3.9 Find the norm of

(a)

$$\begin{pmatrix} 7 & 0 \\ 0 & -9 \end{pmatrix}$$

(b)

$$\begin{pmatrix} 3 & 2 \\ 2 & 0 \end{pmatrix}$$

(c)

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

(d)

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

(e)

$$\begin{pmatrix} 2 & 0 \end{pmatrix}$$

3.10 Prove that if $v \neq 0 \in \mathbb{R}^n$ and if A is an invertible $n \times n$ matrix, then

$$\frac{\|v\|}{\|A^{-1}\|} \leq \|Av\| \leq \|A\| \, \|v\| \, .$$

Lemma 3.6. *If a matrix has small entries then it has small norm. To be precise, suppose that all entries of a $p \times q$ matrix A are smaller in absolute value than some positive number ε. Then $\|A\| < \sqrt{p}\varepsilon$.*

Proof. Take any x with $\|x\| = 1$. Then $y = Ax$ has entries

$$y_i = \sum_j A_{ij} x_j .$$

So

$$\begin{aligned} y_i^2 &\leq \sum_j \varepsilon^2 x_j^2 \\ &= \varepsilon^2 \|x\|^2 . \end{aligned}$$

So

$$\begin{aligned} \|y\|^2 &= \sum_i y_i^2 \\ &\leq \varepsilon^2 p \|x\|^2 . \end{aligned}$$

□

Theorem 3.7. *Suppose that $B \subset \mathbb{R}^n$ is a closed or open ball and $f \colon B \to \mathbb{R}^n$ is a continuously differentiable map. Suppose that $c \geq 0$ is a constant real number and that $\|f'(x)\| < c$ for all x. Then*

$$\|f(s) - f(t)\| \leq c \|s - t\|$$

for all s and t in B.

The proof is identical to that of theorem 3.3 on page 24, using the integration

$$f(t) - f(s) = \int_0^1 f'(ut + (1-u)s)(t-s) \, du.$$

Review problems

3.11
- (a) Prove that the map $f(x) = \arctan \sin \arctan x$ is a contraction map.
- (b) Is $g(x) = \log\left(1 + e^x\right)$. a contraction map? Does it have a fixed point?
- (c) Is $h(x) = \sin x$ a contraction map?

3.12 Suppose that $f \colon \bar{B} \to \mathbb{R}^n$ is a continuous map on a closed ball $\bar{B} \subset \mathbb{R}^n$ and suppose that $\|f'(x) - I\| < \varepsilon$ at every point $x \in \bar{B}$. Imitate the proof of theorem 3.7 on the preceding page to prove that

$$(1-\varepsilon)\,\|s-t\| \le \|f(s)-f(t)\| \le (1+\varepsilon)\,\|s-t\|\,,$$

for every pair of points $s, t \in \bar{B}$.

Chapter 4

The Inverse Function Theorem

We generalize the inverse function theorem from single variable calculus to a very similar theorem in many variable calculus.

Statement and proof

A map $f\colon \text{open } \subset \mathbb{R}^q \to \mathbb{R}^p$ is *continuously differentiable* if f is differentiable and $f'(x)$ is continuous in x (i.e. the matrix entries of $f'(x)$ depend continuously on x).

In one variable calculus, when $y = f(x)$ is a continuously differentiable invertible function,

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)},$$

where $y = f(x)$. There is an inverse function (locally) wherever the right hand side is defined, i.e. away from critical points. Consider the function $y = f(x) = x^2$: This function is not invertible, because each $y > 0$ has two square roots $x = \pm\sqrt{y}$: in any neighborhood of 0. But $f(x)$ is locally invertible near any point x which lies away from 0. There are two local inverse functions defined near each nonzero value: $-\sqrt{y}$ and $\sqrt{y}$. So the inverse, if it exists, might not be unique, but it is “locally unique”.

Theorem 4.1. *Suppose that $U \subset \mathbb{R}^n$ is an open set, that $f\colon U \to \mathbb{R}^n$ is a continuously differentiable map, and that x_0 is a point of U. The matrix $f'(x_0)$ is invertible just when, after perhaps replacing the open set U by a smaller open set containing x_0, f has a differentiable inverse and*

$$\left(f^{-1}\right)'(y) = f'(x)^{-1},$$

if $y = f(x)$.

Roughly: if the linear approximation $f'(x_0)$ is invertible, then locally f is invertible. Note that if f is continuously differentiable, then f^{-1} is too, by the explicit equation for $\left(f^{-1}\right)'$.

Proof. We can clearly replace the function $f(x)$ by the function $f(x - x_0)$ if needed to arrange that $x_0 = 0$, and this won't change the truth of our theorem. Similarly we can replace $f(x)$ by $f(x) - f(0)$ to arrange that $f(0) = 0$. Let

$L = f'(0)$; by the chain rule, we can replace $f(x)$ by $L^{-1}f(x)$, so we can arrange $f'(0) = I$. (We make these replacements just for ease of notation.) After these replacements, $f(x) = x + o(x)$. Pick a value of y near 0. (We will soon see how near it has to be.) Let $F(x) = x + y - f(x)$. The value of y is frozen, so $F(x)$ depends only on x. A fixed point of F, i.e. an x with $F(x) = x$, is precisely a point where $y = f(x)$. We claim that F is a contraction map on some closed ball around 0. The matrix $F'(x) = I - f'(x)$ varies continuously with x and vanishes at $x = 0$. By continuity of $f'(x)$, all the entries of $F'(x)$ are as small as we like if we keep x near enough to 0. By theorem 3.6 on page 26, the norm of $F'(x)$ is as small as we like. For whatever ε we like, say with $0 < \varepsilon < 1$, as long as x stays at most some distance δ from the origin, we can ensure that $\|F'(x)\| < \varepsilon$.

Since $F(0) = y$,

$$\begin{aligned} F(x) - y &= F(x) - F(0), \\ &= \int_0^1 \frac{d}{du} F(ux)\, du, \\ &= \int_0^1 F'(ux) x\, du \end{aligned}$$

so $\|F(x) - y\| < \varepsilon\delta$. We now choose y: take as y any point with $\|y\| \le (1-\varepsilon)\delta$. So if $\|x\| \le \delta$ then $\|F(x)\| \le \delta$. So every $x \in \bar{B}_\delta(0)$ is mapped back to some $F(x) \in \bar{B}_\delta(0)$, i.e. $F\colon \bar{B}_\delta(0) \to \bar{B}_\delta(0)$. By theorem 3.7 on page 26, F is a contraction map on that ball. By theorem 3.4 on page 25, F has a fixed point, say $x \in \bar{B}_\delta(0)$, so $y = f(x)$. Therefore every point $y \in \bar{B}_{\delta(1-\varepsilon)}(0)$ lies in the image of f. By theorem 3.5 on page 25, the point $x = f^{-1}(y)$ depends continuously on the choice of y.

How much does f squish points together? Not much: by problem 3.12 on page 27,

$$\|f(s) - f(t)\| \ge (1-\varepsilon)\|s - t\|,$$

for all $s, t \in \bar{B}_\delta(0)$. Therefore, if $y = f(x)$,

$$\begin{aligned} \frac{\left\|f^{-1}(y) - y\right\|}{\|y\|} &= \frac{\|x - f(x)\|}{\|f(x)\|}, \\ &= \frac{o(x)}{\|f(x)\|}, \\ &\le \frac{o(x)}{(1-\varepsilon)\|x\|}, \\ &\le \frac{1}{1-\varepsilon}\frac{o(x)}{\|x\|}, \\ &\to 0, \end{aligned}$$

as $x \to 0$, i.e. as $y \to 0$. Therefore $f^{-1}(y) = y + o(y)$, so $\left(f^{-1}\right)'(0) = I = f'(0)^{-1}$. □

Review problems

4.1 Prove that the map

$$(u, v) = f(x, y) = \left(x^2 - y^2, 2xy\right)$$

has a local inverse defined near each point $(x, y) \neq (0, 0)$.

4.2 Draw a picture to explain the inverse function theorem in one variable calculus.

4.3 Draw pictures to explain some of the different things that can happen when $f'(x_0) = 0$ in one variable calculus.

4.4 Give an example to show that if $f'(x_0)$ is *not* invertible, then $f(x)$ might have no inverse near $x = x_0$. Give another example to show that if $f'(x_0)$ is *not* invertible, then $f(x)$ could still have an inverse near $x = x_0$, and even a continuous inverse. Prove that if $f'(x_0)$ is *not* invertible, then $f(x)$ does *not* have a continuously differerentiable inverse.

4.5 Suppose that $f\colon U \to \mathbb{R}^n$ is a continuously differentiable map, $U \subset \mathbb{R}^n$ is an open set, and $f'(x)$ is an invertible matrix for every $x \in U$. Prove that $f(U) \subset \mathbb{R}^n$ is an open set.

4.6* We want to see that the roots of a one variable polynomial depend continuously on the coefficients. For any numbers $r = (r_1, r_2, \dots, r_n) \in \mathbb{R}^n$, let p_r be the polynomial

$$p_r(x) = (x - r_1)(x - r_2)\dots(x - r_n),$$

with roots $r_1, r_2, \dots, r_n$. We can write $p_r(x)$ in terms of its coefficients, say

$$p_r(x) = x^n + \sum_{j=1}^{n} a_j x^{j-1}.$$

These $a_1, a_2, \dots, a_n$ coefficients depend on the choice of r, say as a map $a = a(r)$, so $a\colon \mathbb{R}^n \to \mathbb{R}^n$.

(a) Prove that

$$\frac{\partial p_r(x)}{\partial r_i} = -\prod_{j \neq i} (x - r_j).$$

(b) Explain why this says that

$$\sum \frac{\partial a_j}{\partial r_i} x^{j-1} = -\prod_{j \neq i} (x - r_j).$$

(c) Suppose that $a'(r)s = 0$, for some vector $s \in \mathbb{R}^n$. Prove that

$$\sum_i \frac{\partial p_r(x)}{\partial r_i} s_i = 0.$$

Use this to show that

$$0 = -\sum_i \prod_{j \neq i} (x - r_j)\, s_i.$$

for all x.

(d) Suppose that all of the numbers $r_1, r_2, \dots, r_n$ are distinct. Plug in $x = r_k$ to the last result for some k to prove that if $a'(r)s = 0$ then $s_k = 0$. Conclude that $a'(r)$ has kernel 0 at any point $r \in \mathbb{R}^n$ with all of the numbers $r_1, r_2, \dots, r_n$ distinct.

(e) Explain why this says that if a polynomial of degree n has n distinct roots, then every polynomial whose coefficients are close enough has n distinct roots, and these roots vary as continuously differentiable functions of the coefficients.

4.7* Continuity of eigenvalues and eigenvectors: suppose that $A(t)$ is an $n \times n$ matrix depending as a continuously differentiable function on a real variable t. Suppose that at $t = 0$, the matrix $A_0 = A(0)$ has an eigenvalue λ_0 and that the generalized λ_0-eigenspace is 1-dimensional containing a λ_0-eigenvector x with $x_1 = 1$. Prove that, for all t close to 0, there is a unique continuously differentiable choice of eigenvalue $\lambda(t)$ of $A(t)$ so that $\lambda(0) = \lambda_0$, and there is a unique continuously differentiable eigenvector $x(t)$ with eigenvalue $\lambda(t)$ and with $x_1(t) = 1$. Hint: let $F(x, \lambda, t) = (A(t)x - \lambda x, x_1 - 1, t)$, $F \colon \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$. If $F'(x, \lambda, t)$ is invertible, then F is locally invertible. Check that

$$F'(x, \lambda, t)(y, \mu, s) = (A(t)y + A'(t)sx - \lambda y - \mu x, y_1, s).$$

4.8 Suppose that $f \colon \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function of one variable, and $f'(x) \neq 0$ for all $x \in \mathbb{R}$. Prove that f is 1-1.

4.9 Let $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ be the map

$$(u, v) = f(x, y) = (e^x \cos y, e^x \sin y).$$

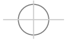
Prove that $f'(x, y)$ is invertible for all (x, y), but f is not 1-1.

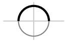
Chapter 5

The Implicit Function Theorem

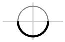
We see how to check that the set of solutions of a system of equations form a smooth object.

Statement and proof

A circle $x^2 + y^2 = 1$  is locally the graph of a function



$y = \sqrt{1 - x^2}$

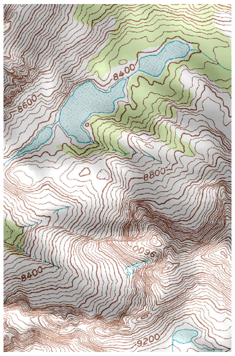


$y = -\sqrt{1 - x^2}$

At the left and right ends, both functions have nasty derivatives: the circle slopes up vertically.

For any map, say F, and any constant c, the *c-level set* of F is the set $\{\, x \mid F(x) = c \,\}$. Think of a geographical map of mountains, F being the height of each point—the level sets are the contours.

If $F(x, y)$ is a function of two variables (for example, take the function $F(x, y) = x^2 + y^2$ to recover our example above) then anywhere where $\frac{\partial F}{\partial y} \neq 0$, we will see that each level set of F is locally the graph of a function $y = f(x)$.



Level sets

Theorem 5.1 (Implicit function theorem)**.** *Suppose that $U \subset \mathbb{R}^{p+q}$ is an open set and $F \colon U \to \mathbb{R}^q$ is a continuously differentiable function. Write each point of U as a pair (x, y) with $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$. Write $\frac{\partial F}{\partial y}$ to mean the matrix*

$$\frac{\partial F}{\partial y} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} & \dots & \frac{\partial F_1}{\partial y_q} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} & \dots & \frac{\partial F_2}{\partial y_q} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial F_q}{\partial y_1} & \frac{\partial F_q}{\partial y_2} & \dots & \frac{\partial F_q}{\partial y_q} \end{pmatrix}.$$

If $\frac{\partial F}{\partial y}$ is an invertible matrix at a point $(x, y) = (x_0, y_0)$ then, after perhaps replacing U by a smaller open set containing (x_0, y_0), every level set of F is the graph $y = f(x)$ of a continuously differentiable map f defined on some open set in $\mathbb{R}^q$.

Before the proof, an example. To draw a circle, take $F(x,y) = x^2 + y^2$, $U = \mathbb{R}^2$, and the circle is the set

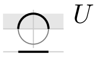
$$\{\, (x,y) \in U \mid F(x,y) = 1 \,\}.$$

Try the point $(x_0, y_0) = (0,1)$. Then

$$\frac{\partial F}{\partial y} = 2y = 2.$$

(This is a 1×1 matrix, i.e. a number.) So near this point, we can solve for y as a function of x. We can replace U by letting

$$U = \{\, (x,y) \mid y > 0 \,\} \subset \mathbb{R}^2.$$



The only function $y = f(x)$ whose graph is given by this portion of the circle is $f(x) = \sqrt{1 - x^2}$.

Proof. Let $G(x,y) = (x, F(x,y))$ for any $(x,y) \in U$. Check that

$$G'(x,y) = \begin{pmatrix} I & 0 \\ \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{pmatrix}$$

is invertible near $(x,y) = (x_0, y_0)$. By the inverse function theorem, G has an inverse on some open set; shrink U if needed to assume that G is invertible.

If $(X,Y) = G(x,y) = (x, F(x,y))$, then $X = x$, so inverting, if $(x,y) = G^{-1}(X,Y)$, then $x = X$. Therefore

$$G^{-1}(x,y) = (x, g(x,y)),$$

for some continuously differentiable function $g(x,y)$. Let $f(x) = g(x,0)$. If $(X,Y) = G(x,y)$, then $Y = F(x,y)$, so $F(x,y) = 0$ just when $Y = 0$ which happens just when

$$\begin{aligned} (x,y) &= G^{-1}(X,0), \\ &= (X, g(X,0)), \\ &= (x, f(x)). \end{aligned}$$

In other words, $F(x,y) = 0$ just on the graph of $y = f(x)$. □

The rank theorem

This section can be omitted without loss of continuity.

A very slightly stronger implicit function theorem: we can see when a map looks locally like a linear map.

Theorem 5.2 (The rank theorem)**.** *Suppose that $f\colon U \subset \mathbb{R}^p \to \mathbb{R}^q$ is a continuously differentiable map from an open set $U \subset \mathbb{R}^p$, and that $f'(x)$ has the same rank at each point $x \in U$. Then f is locally identified by diffeomorphisms with a linear map. In other words, for each point $x_0 \in U$, there is a diffeomorphism $F\colon U_0 \to U_1$ of a neighborhood U_0 of x_0 and a diffeomorphism $G\colon W_0 \to W_1$ of a neighborhood W_0 of $f(x_0)$, so that $G \circ f \circ F^{-1}$ is a linear map.*

Proof. The matrix $f'(x)$ has rank k. Write the map f as $y = f(x)$ and relabel the x variables and the y variables to arrange that the $k \times k$ upper left corner of f is an invertible matrix. Now instead write the map f as $(u, v) = f(x, y)$, where in these new variables, $u \in \mathbb{R}^k$ and $v \in \mathbb{R}^{q-k}$ and $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^{p-k}$, say $(u, v) = f(x, y)$. So $\frac{\partial u}{\partial x}$ is invertible. Let $F(x, y) = (u(x, y), y)$ so that $f\left(F^{-1}(u, y)\right) = (u, v)$ say. If we write this as $f\left(F^{-1}(u, y)\right) = (u, v(u, y))$ for simplicity, allowing ourselves the freedom now to think of u as a variable, not a map, then we differentiate

$$\left(f \circ F^{-1}\right)'(u, y) = \begin{pmatrix} I_{k\times k} & 0 \\ \frac{\partial v}{\partial u} & \frac{\partial v}{\partial y} \end{pmatrix}.$$

The rank of this matrix must be k, because it is the product $(f' \circ F)F'$ where F' is invertible. But there is already rank k in the upper left. If there were any nonzero entries in the lower right, they would increase the rank. So $\frac{\partial v}{\partial y} = 0$, i.e. $f\left(F^{-1}(u, y)\right) = (u, v(u))$ for some map $v(u)$. Let $G(u, v) = (u, v - v(u))$. ☐

Review problems

5.1 Let

$$F(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 - 1 \\ \cos x + \sin y - 1 \end{pmatrix}.$$

Prove that we can solve the equation $F = 0$ for y and z as continuously differentiable functions of x near $(0, 0, 1)$. Can you draw a picture? Can you solve $F = 0$ for x and y as continuously differentiable functions of z near $(0, 0, 1)$?

5.2 Prove that there is no 1-1 continuously differentiable map $f\colon \mathbb{R}^2 \to \mathbb{R}$.

5.3 Consider the curve in $\mathbb{R}^3$ consisting of the intersection of the paraboloid $z = x^2 + y^2$ and the cylinder $x^2 + z^2 = 1$. Near which points of this curve does the implicit function theorem say that we can write the curve as the graph of continuously differentiable functions $y = y(x), z = z(x)$?

5.4* Consider the cubic curve $y^2 = x(x - \lambda)(x - 1)$ in the (x, y)-plane. Prove that y is locally a continuously differentiable function of x except near the points $x = 0, x = \lambda$ and $x = 1$. Prove that the inflection points of the cubic curve (where $\frac{d^2y}{dx^2} = 0$) have coordinates (x, y) given locally by continuously differentiable functions of λ as long as $(x, y) \neq (0, 0)$ and either $(x, y) \neq (1, 0)$ or $\lambda \neq 0, \frac{1}{2}, 1$.

5.5 Consider the sphere $x^2 + y^2 + z^2 = 1$ and the cylinder $(x - a)^2 + y^2 = r^2$ where $r > 0$ and $a \geq 0$. Let M be the intersection of the sphere and cylinder. For what values of a and r and near what points $(x, y, z) \in M$ does the implicit function theorem ensure that M is locally a graph $y = y(x), z = z(x)$? Draw pictures of what the sphere and cylinder look like at all these a and r values, near each of these (x, y, z) points, where the implicit function theorem doesn't apply. Hint: let

$$F(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 - 1 \\ (x - a)^2 + y^2 - r^2 \end{pmatrix}.$$

5.6*

(a) Prove that the product of any two upper triangular square matrices is upper triangular. Hint: induction, and write any upper triangular matrix A as

$$A = \begin{pmatrix} B & c \\ 0 & d \end{pmatrix}$$

where B is upper triangular and d is a number.

(b) Use the same trick to prove that the inverse of any invertible upper triangular matrix is upper triangular.

(c) It is well known that every invertible matrix A can be written as $A = QR$ where Q is an orthogonal matrix and R is an upper triangular matrix whose diagonal entries are positive. Use the inverse function theorem to prove that Q and R are continuously differentiable functions of A. Hint: Write $\mathbb{R}^{n^2}$ for the set of all $n \times n$ matrices, $\mathbb{R}^{n(n+1)/2}$ for the set of all symmetric matrices, and $\mathbb{R}^{n(n-1)/2}$ for the set of all antisymmetric matrices. For any matrix R, let $L(R)$ be the antisymmetric matrix which has the same entries as R below the diagonal. Let $F(Q, R) = (QR, QQ^t - I, L(R))$,

$$F \colon \mathbb{R}^{n^2} \times \mathbb{R}^{n^2} \to \mathbb{R}^{n^2} \times \mathbb{R}^{n(n+1)/2} \times \mathbb{R}^{n(n-1)/2}.$$

Show that

$$F'(Q_0, R_0)(Q_1, R_1) = \left(Q_0 R_1 + Q_1 R_0, Q_0 {Q_1}^t + Q_1 {Q_0}^t, L(R_1)\right),$$

for any matrices Q_0, Q_1, R_0, R_1.

5.7* It is well known that there is an open set of invertible matrices A that can be written as $A = LU$ where L is a lower triangular matrix with 1's down the diagonal, and U is an invertible upper triangular matrix, and L and U are unique. Prove that L and U are continuously differentiable functions of A. Hint: much as in the previous problem, let V_+ be the vector space of upper triangular matrices, V_- the vector space of lower triangular matrices, and map

$$F \colon V_- \times V_+ \to \mathbb{R}^{n^2}, F(L, U) = LU.$$

5.8* The polar decomposition: it is well known that every invertible matrix A can be written uniquely as $A = KP$ where K is an orthogonal matrix and P is an invertible symmetric matrix all of whose eigenvalues are positive. Prove that K and P are continuously differentiable functions of A. Hint: by the spectral theorem, for any fixed choice of P, you can always pick an orthonormal basis in which P is diagonal, so at just the right point in your proof you can assume that P is diagonal.

Notes

1. The condition $\frac{\partial F}{\partial y} \neq 0$ says that it looks like there is a y in the equation $F(x, y) = 0$. If $\frac{\partial F}{\partial y} = 0$, it looks (as far as first derivatives are concerned) as if there is no y.
2. The linear approximation to an equation $F(x, y) = 0$ near a point (x_0, y_0) is the equation
$$\frac{\partial F}{\partial x}(x_0, y_0)\, x + \frac{\partial F}{\partial y}(x_0, y_0)\, y = 0.$$
The Implicit Function Theorem says that you can solve for y in the linear approximation just when you can solve for y in the original fully nonlinear equation.

Chapter 6

Manifolds

We define manifolds; intuitively these are just smooth objects.

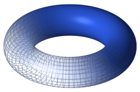
What manifolds are

A circle in the plane is the same sort of thing as a circle drawn in 3-dimensional space. But in what sense? A *diffeomorphism* is a continuously differentiable map $f\colon X \to Y$ between two sets $X \subset \mathbb{R}^p$ and $Y \subset \mathbb{R}^q$ whose inverse is also a continuously differentiable map $f^{-1}\colon Y \to X$; the sets X and Y are *diffeomorphic*. A set X is *locally diffeomorphic* to a set Y if every point of X lies in an set diffeomorphic to some open set in Y.

6.1 Give an example of sets X and Y with X locally diffeomorphic to Y, but Y not locally diffeomorphic to X.

6.2 Prove that if $f\colon \text{open } \subset \mathbb{R}^p \to \text{open } \subset \mathbb{R}^q$ is a diffeomorphism then $p = q$. Hint: look at f'.

A manifold means a "smooth object", roughly speaking. A *p-dimension manifold* is a set M locally diffeomorphic to $\mathbb{R}^p$. For example, a 1-dimensional manifold in $\mathbb{R}^2$ is a curve in the plane, a 1-dimensional manifold in $\mathbb{R}^3$ is a curve in space, a 2-dimensional manifold in $\mathbb{R}^3$ is a surface in space, and an n-dimensional manifold in $\mathbb{R}^n$ is just precisely an open set in $\mathbb{R}^n$.



A typical manifold

6.3 Suppose that $M \subset \mathbb{R}^n$ is a nonempty manifold of dimension p. Prove that M is not a manifold of any other dimension $q \neq p$.

We want a simple way to write down manifolds. A *chart* is a continuously differentiable map $f\colon U \text{ open } \subset \mathbb{R}^p \to \mathbb{R}^n$ so that there is an open set $W \subset \mathbb{R}^n$ and a continuously differentiable map $g\colon W \to U$ with $g(f(x)) = x$ for every $x \in U$. Roughly speaking, f has a "left inverse" g. For example, the map

$$f(s,t) = (\cos s, \sin s, t)\,,\ -\frac{\pi}{2} < s < \frac{\pi}{2},$$

has a left inverse

$$g(x,y,z) = (\arctan(y/x), z)\,,$$

because by elementary trigonometry, $g(f(s,t)) = (s,t)$. A manifold is locally diffeomorphic to $\mathbb{R}^p$; if $f\colon \text{open } \subset \mathbb{R}^p \to \text{open } \subset M$ is one such diffeomorphism, the map f is also a chart. Conversely, roughly speaking, a manifold is just a set which is locally the image of a chart: precisely, a set M is a manifold just when every point $p \in M$ lies in an open set W so that $M \cap W$ is the image of a chart. For example, the circle is a manifold because the circle is locally the image of the chart

$$f(\theta) = (\cos\theta, \sin\theta),$$

say for some interval of values of θ on which we have some left inverse given by a trigonometric function $g(x, y) = \arctan \frac{y}{x}$ or $g(x, y) = \arccos x$ or $g(x, y) = \arcsin y$ etc.

6.4 Prove that the following are manifolds, by exhibiting charts and left inverses for those charts: (1) a circle, (2) a parabola, (3) a sphere, (4) the graph of a continuously differentiable function $y = h(x)$.

This is not the easiest way to write down manifolds; it is easier to use the implicit function theorem.

Theorem 6.1 (Implicit function theorem)**.** *Suppose that $f \colon U \subset \mathbb{R}^{p+q} \to \mathbb{R}^q$ is a continuously differentiable map from an open set $U \subset \mathbb{R}^{p+q}$, and for each point $x \in U$, $f'(x)$ is a surjective linear map. Then every level set of f is a manifold.*

For example, a circle is a manifold; take $f(x, y) = x^2 + y^2$, $U = \mathbb{R}^2 - \{(0, 0)\}$.

Proof. After we perhaps restrict to a smaller open set, and perhaps permute the order of the coordinates $x_1, x_2, \dots, x_n$, we can arrange that

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_q}$$

form a basis. Write x as $x = (s, t)$, $s \in \mathbb{R}^q$ and $t \in \mathbb{R}^p$. By the implicit function theorem, we can arrange that each level set of f is the graph $t = g(s)$ of some continuously differentiable function $g \colon \text{open} \subset \mathbb{R}^q \to \mathbb{R}^p$. Our chart is $h(s) = (s, g(s))$, with left inverse $k(s, t) = s$. ☐

6.5 Use the implicit function theorem to prove that the intersection of the paraboloid $z = x^2 + y^2$ with the paraboloid $y = x^2 + z^2$ is a manifold in $\mathbb{R}^3$. What is the dimension of this manifold?

6.6 Consider the sphere $x^2 + y^2 + z^2 = 1$ and the cylinder $(x - a)^2 + y^2 = r^2$ where $r > 0$ and $a \geq 0$. Let M be the intersection of the sphere and cylinder. For what values of a and r and near what points $(x, y, z) \in M$ does the implicit function theorem ensure that M is a manifold? Draw pictures of what the sphere and cylinder look like at all these a and r values, near each of these (x, y, z) points, where the implicit function theorem doesn't apply? Hint: let

$$F(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 - 1 \\ (x - a)^2 + y^2 - r^2 \end{pmatrix}.$$

6.7 Use the implicit function theorem to prove that the intersection of the paraboloid $z = x^2 + y^2$ with the cylinder $x^2 + z^2 = 4$ is a manifold in $\mathbb{R}^3$. What is the dimension of the manifold?

6.8 Prove that the graph $y = f(x)$ of a function $f\colon \text{open } \subset \mathbb{R}^p \to \mathbb{R}^q$ is a p-dimensional manifold in $\mathbb{R}^{p+q}$ just when f is continuously differentiable.

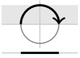
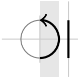
6.9* An easier definition of manifold to imagine: prove that a set $M \subset \mathbb{R}^2$ is a 1-dimensional manifold precisely when each point of M lies in an open set $W \subset \mathbb{R}^2$ in which $M \cap W$ is precisely the graph of one of the variables set to be a function of the other variable. (For example, the circle $x^2 + y^2 = 1$ has top half $y = \sqrt{1 - x^2}$, bottom half $y = -\sqrt{1 - x^2}$, right half $x = \sqrt{1 - y^2}$, and left half $x = -\sqrt{1 - y^2}$.) Generalize this to p-dimensional manifolds in $\mathbb{R}^{p+q}$.

6.10* Prove that the union of two lines in the plane is a manifold if and only if the lines are parallel. Hint: a manifold of dimension 1 could locally cut into two pieces if you cut out a point.

6.11 Prove that $[0, \infty)$ is not a manifold.

6.12* Let $\mathrm{O}(n)$ be the set of orthogonal $n \times n$ matrices. Prove that $M \subset \mathbb{R}^{n^2}$ is a manifold of dimension $n(n-1)/2$.

Suppose that h_1 and h_2 are two charts on a p-dimensional manifold M, and their images overlap in M. Then the map $H = h_2^{-1} \circ h_1$ (wherever it is defined) is called the *transition map* between the two charts. The transition map is a diffeomorphism between two open sets in $\mathbb{R}^p$.

For example, let M be the unit circle. We can make a chart along the top  by keeping track of the x coordinate of each point. We can make a chart along the right side  by keeping track of the y coordinate of each point. The transition from the first of these charts to the second one is just the map taking the x coordinate to the y coordinate, defined where both are defined, i.e. on $0 < x < 1$ and $0 < y < 1$, so the transition map is

$$H(x) = y = \sqrt{1 - x^2}.$$

6.13 The unit circle has charts

$$h_1(s) = (\cos s, \sin s) \in \mathbb{R}^2,$$

for $-\pi < s < \pi$, and

$$h_2(t) = \left(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right)$$

for $t \in \mathbb{R}$. Compute the transition map $H = h_2^{-1} \circ h_1$.

The rank theorem

The kernel of any linear map is a linear subspace, so a We can restate the rank theorem to take advantage of this.

Theorem 6.2 (The rank theorem)**.** *Suppose that $f \colon U \subset \mathbb{R}^p \to \mathbb{R}^q$ is a continuously differentiable map from an open set $U \subset \mathbb{R}^p$, and that $f'(x)$ has the same rank at each point $x \in U$. Then every level set of f is a manifold of dimension $p - k$.*

For example, on the space $\mathbb{R}^{n^2}$ of $n \times n$ matrices, take the map $f \colon \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, $f(A) = A^t A$. The kernel of $f'(A)$ is precisely the set of matrices B so that $A^t B + B^t A = 0$. Let $U \subset \mathbb{R}^{n^2}$ be the set of invertible matrices, an open set. If A is invertible, let $C = A^t B$ and recover B as $B = \left(A^t\right)^{-1} C$. Then B lies in the kernel of $f'(A)$ just when $C + C^t = 0$, i.e. C is antisymmetric. The subspace of antisymmetric matrices has dimension $n(n-1)/2$. So $f'(A)$ has kernel of dimension $n(n-1)/2$, and so rank $n^2 - n(n-1)/2 = n(n+1)/2$. By the rank theorem, for any invertible matrix S, the level set $M_S = \{\, A \mid f(A) = S \,\}$ is a manifold of dimension $n(n-1)/2$. The manifold consists precisely of all invertible matrices A so that $A^t A = S$. Roughly speaking, M_S is the set of square roots of the symmetric matrix S. The spectral theorem shows that M_S is not empty just when all of the eigenvalues of S are positive.

Manifolds with boundary

The n-dimensional *left half space* is the set $(-\infty, 0] \times \mathbb{R}^{n-1}$, consisting of those $x \in \mathbb{R}^n$ for which $x_1 \le 0$. Its *boundary* is the set of $x \in \mathbb{R}^n$ with $x_1 = 0$. A *n-dimensional manifold with boundary* is a set locally diffeomorphic to the n-dimensional left half space. Consider the n-dimensional unit box

$$X = [0,1] \times [0,1] \times \dots \times [0,1] \subset \mathbb{R}^n.$$

The *boundary* of X is the set of points $x \in X$ at which $x_j = 0$ or $x_j = 1$ for some j. A *n-dimensional manifold with boundary and corners* is a set locally diffeomorphic to the unit n-dimensional box. The inverse map f^{-1} of the local choice of diffeomorphism is called a *chart*. The *boundary* is the set of points identified by some chart with boundary points of the box. Careful: the term "boundary" is also used in mathematics with a different meaning (but not in these notes). The set of boundary points of M is denoted ∂M. For example, the boundary of is .

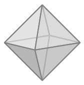
6.14 Prove that the boundary of a manifold with boundary is a manifold.

6.15 Give an example a manifold with boundary and corners whose boundary is not a manifold with boundary and corners.

6.16 Suppose that M is a 2-dimensional manifold with boundary and corners. Do the angles at the corners have to be right angles?

6.17 Prove that a hollow cylinder in $\mathbb{R}^3$ is a manifold with boundary.

6.18 Prove that a solid cylinder in $\mathbb{R}^3$ is a manifold with boundary and corners.



6.19* Is an octahedron a manifold with boundary and corners?

6.20* Is a cone a manifold with boundary and corners?

6.21 Suppose that $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}$ is a continuously differentiable map, and that $f'(x) \neq 0$ for every $x \in U$. Prove that the set

$$M = \{\, x \in U \mid f(x) \leq 0 \,\} \subset \mathbb{R}^n$$

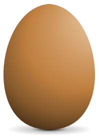
is a manifold with boundary.

6.22 Suppose that $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}^n$ is a continuously differentiable map, and that $\det f'(x) \neq 0$ for every $x \in U$. Prove that the set

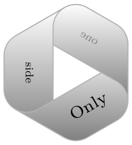
$$M = \{\, x \in U \mid f_1(x) \geq 0 \text{ and } f_2(x) \geq 0 \text{ and } \dots \text{ and } f_n(x) \geq 0 \,\} \subset \mathbb{R}^n$$

is a manifold with boundary and corners.

Orientation



A surface from outside



A Möbius strip

You can travel along a curve either of two ways: ↻ or ↺. A choice of one of these directions is called an *orientation.* A surface in $\mathbb{R}^3$ has an inside and an outside.

Actually, that isn't quite true: the Möbius strip has only one side.

If a surface has two sides, a choice of side is called an *orientation.* A 0-dimensional manifold is a collection of isolated points. By definition, an orientation of a 0-dimensional manifold is just a choice of ± sign for each point.

A differentiable map $f\colon \text{open} \subset \mathbb{R}^n \to \mathbb{R}^n$ is *orientation preserving* if $\det f'(x) > 0$ at every point x where f is defined. For example, if $n = 1$ this says that $f(x)$ travels along $\mathbb{R}$ in the same direction as x. Recall that every diffeomorphism f is approximately linear, approximated by $f'(x)$. If $n = 2$, this says that f takes small curves in the plane travelling counterclockwise to small curves in the plane also travelling counterclockwise. If $n = 3$, this says that f takes a small sculpture of your right hand into something that looks like a slightly deformed right hand, *not* a slightly deformed left hand. Two charts are *compatibly oriented* if their transition map is orientation preserving. Returning to our example of the unit circle, we had two charts with transition map

$$H(x) = y = \sqrt{1 - x^2}.$$

So

$$H'(x) = \frac{-x}{\sqrt{1 - x^2}} < 0.$$

Therefore these two charts are not compatibly oriented: the two charts travel in opposite directions around the circle.

6.23 Consider the charts

$$f(s) = (\cos s, \sin s) \in \mathbb{R}^2,$$

for $-\pi < s < \pi$, and

$$g(t) = \left(\frac{2t}{t^2+1}, \frac{t^2-1}{t^2+1}\right)$$

for $t \in \mathbb{R}$. Are these charts on the circle compatibly oriented?

6.24 Consider the *cylindrical coordinate chart*

$$(x, y, z) = f(r, \theta, Z) = (r \cos\theta, r \sin\theta, Z)$$

as a map

$$f \colon (0, \infty) \times (-\pi, \pi) \times \mathbb{R} \to \mathbb{R}^3,$$

and the *spherical coordinate chart*

$$(x, y, z) = g(R, \alpha, \beta) = (R \cos\alpha \cos\beta, R \sin\alpha \cos\beta, R \sin\beta)$$

as a map

$$g \colon (0, \infty) \times (-\pi, \pi) \times (-\pi/2, \pi/2) \to \mathbb{R}^3.$$

Compute the transition map $h = g^{-1} \circ f$ and check whether these two charts on $\mathbb{R}^3$ are orientation compatible near $(x, y, z) = (1, 0, 0)$.

6.25* Suppose that A is a square matrix. Prove that $\det A > 0$ just when there is a continuous map associated to each real number t a matrix $B(t)$ so that $B(0) = A$ and $B(1) = I$. Hint: use Gram–Schmidt orthogonalization applied to the columns of A to create a matrix Q. Prove that $A = QR$ where R is upper triangular. It is easy to deform R, and a little harder to deform Q. Hint: split $\mathbb{R}^n$ into 2-dimensional Q-invariant subspaces.

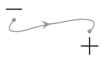
An *orientation* of a manifold $M \subset \mathbb{R}^n$ of positive dimension is a collection of charts so that

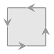
1. any two of the charts are compatibly oriented and
2. a chart belongs to the collection just when it is compatibly oriented with any element of the collection and
3. every point of M lies in the image of a chart.

These charts are called the *positively oriented charts* of the orientation of the manifold.

6.26 If $M \subset \mathbb{R}^p$ and $N \subset \mathbb{R}^q$ are oriented manifolds, how could we orient $M \times N \subset \mathbb{R}^{p+q}$?

Orienting the boundary

We orient the boundary of a curve by putting a + sign on any point the curve travels to, and a − sign on any point the curve travels from: 

We orient the boundary of any open set in the plane by the right hand rule:  as you walk along the boundary, your right hand sticks out away from the set.

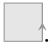
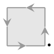
Let's generalize the right hand rule to n dimensions. (Don't worry: you only need to remember the right hand rule, and the fact that there is some way to generalize it. You don't have to remember the details in higher dimensions.) Orient the left half space $(-\infty, 0] \times \mathbb{R}^{n-1}$ by seeing it as a subset of $\mathbb{R}^n$; orient its boundary by the chart $\mathbb{R}^{n-1} \to \{0\} \times \mathbb{R}^{n-1}$:

$$(x_2, u_3, \dots, x_n) \mapsto (0, x_2, x_3, \dots, x_n)\,.$$

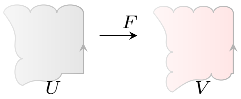
Orient the boundary of any oriented manifold with boundary by using charts to locally identify with the left half space. We won't ever want to orient near corner points, so we ignore those.

Let's see how this gives us the right hand rule. Look at the square $X = [0,1] \times [0,1]$. Near the right end of the square, we have a positively oriented chart

$$(x, y) \in \text{left half space} \mapsto (x+1, y) \in X.$$

The induced chart on the right end wall of the square is given by setting $x = 0$, so $y \mapsto (1, y)$, going up the vertical direction: . Rotations are orientation preserving, so we can rotate the picture to see how the other sides of the square are oriented: .

Lemma 6.3. *If $F\colon U \to V$ is a positively oriented diffeomorphism between two open subsets of the left half space, then the induced map on each of the sets $\partial U \to \partial V$ is orientation preserving. Consequently, along each wall of an oriented manifold with boundary and corners an induced orientation is consistently defined.*



Proof. For simplicity, we assume that U and V are open subsets of the left half space in the plane. The proof in general is similar. We write our map as

$$(s, t) = F(x, y).$$

Then U lies in the set $x \le 0$, and V lies in $s \le 0$.

A point $(x, y) \in U$ lies in the boundary just when there is a function f with $f'(x, y) \neq 0$ but (x, y) a local maximum for f: for example the function $f(x, y) = x$. Since the existence of such a function is a property invariant under diffeomorphism, any diffeomorphism takes boundary points to boundary points.

The map along the wall is just $(0, t(0, y))$, i.e. $s(0, y) = 0$. Therefore $\frac{\partial s}{\partial y} = 0$ along the wall. Along the wall, $\frac{\partial s}{\partial x} \geq 0$; otherwise moving x to the left a little would make s positive. So now

$$F'(0, y) = \begin{pmatrix} \frac{\partial s}{\partial x} & 0 \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{pmatrix}$$

and $0 < \det F'(0, y) = \frac{\partial s}{\partial x} \frac{\partial t}{\partial y}$, so $\frac{\partial s}{\partial x} > 0$ and $\frac{\partial t}{\partial y} > 0$. $\square$

Review problems

6.27 Pick positive constants $r_1 > r_0$. The *torus* is the image of the map

$$F \colon \mathbb{R}^2 \to \mathbb{R}^3, (\alpha, \beta) \mapsto (r_1 + r_0 \cos\alpha \cos\beta, r_1 + r_0 \cos\alpha \sin\beta, r_0 \sin\alpha).$$

Prove that the torus is a 2-dimensional manifold in $\mathbb{R}^3$.

6.28 Draw arrows to orient the boundary of an annulus by the right hand rule.

6.29 Find an example of a system of algebraic equations, say $F(x) = 0$, whose set of solutions M is a manifold, but so that $F'(x)$ does not have full rank at any point of M.

6.30* Denote by $\mathrm{SL}(n, \mathbb{R})$ the set of $n \times n$ matrices of determinant 1. Prove that $\mathrm{SL}(n, \mathbb{R}) \subset \mathbb{R}^{n^2}$ is a manifold of dimension $n^2 - 1$.

6.31* Denote by $\mathrm{SO}(n)$ the set of $n \times n$ rotation matrices (in other words: orthogonal matrices of determinant 1). Prove that $\mathrm{SO}(n) \subset \mathbb{R}^{n^2}$ is a manifold of dimension $n(n-1)/2$.

Notes

1. The objects we refer to as manifolds are more precisely called *embedded submanifolds* of Euclidean space $\mathbb{R}^n$. There is a more abstract definition of *manifold* in the literature, and a deep proof that every manifold in that abstract sense actually arises as an embedded submanifold of $\mathbb{R}^n$.
2. We often give proofs for manifolds and let the reader generalize the proofs for manifolds with boundary or corners, out of laziness, when the proofs turn out essentially the same.
3. A 1-dimensional manifold is called a *curve*, while a 2-dimensional manifold is called a *surface*. Careful: we also call a continuously differentiable map $x \colon I \to \mathbb{R}^n$ a curve, if I is an interval of $\mathbb{R}$. These two uses of the word *curve* are *not* equivalent.
4. Another way to think about orientations: say that a basis $v_1, v_2, \dots, v_n \in \mathbb{R}^n$ is *positively oriented* if $v_j = A e_j$ for some matrix A with $\det A > 0$. If $v_2, v_3, \dots, v_n$ lie tangent to the boundary of an n-dimensional box, and v_1

“sticks out” of the box, then we declare that $v_2, v_3, \ldots, v_n$ is a positively oriented basis for that wall. We won’t try to follow this approach, or make this precise, but it might help to clarify the right hand rule.

Chapter 7

Integration

We describe the properties of integrals in several variables. The proofs are the same as in your courses last year.

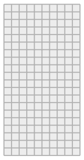
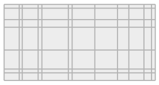
Definition

The *length* of an interval $[a, b]$ is $|b - a|$. The *volume* of a box

$$X = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

is the product of the lengths of the intervals

$$[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n].$$

A grid is something like this:  or this: 

7.1 Give a precise definition of a grid in $\mathbb{R}^n$. Hint: start with $\mathbb{R}$.

The *mesh size* of the grid is the smallest width or height or depth, etc. of any of the small boxes in the grid.

7.2 Give a precise definition of mesh size. Hint: start with $\mathbb{R}$.

We assume that last year you learned about integration, particularly the definition via Riemann sums. Given a box $X \subset \mathbb{R}^n$ and a continuous function f on X and a grid covering X, say made of boxes $X_1, X_2, \dots, X_p$, a *Riemann sum* of f on that grid is a sum

$$\sum_i \operatorname{Vol}(X_i) f(x_i),$$

where $x_i \in X_i$ are some points, which can be chosen arbitrarily. Note: in this sum, we extend f to be 0 outside of X. By the same proof as in one variable calculus, there is a unique number, denoted

$$\int_X f$$

and called the *integral* of f over X so that, if we make a small enough mesh size, the Riemann sum is as close as we want to that number. If f has compact support, we can just let $f = 0$ every in $\mathbb{R}^n$ away from its support and write $\int f$ to mean $\int_X f$ for any large enough box X.

An integral in 2 variables

$$\int f(x, y)\, dx\, dy$$

can be carried out as an iterated one variable integral:

$$\int f(x, y)\, dx\, dy = \int \left(\int f(x, y)\, dy \right) dx,$$

which is clear from summing up the Riemann sums first up a column of the grid, and then summing across. The same idea works in any number of variables: you can integrate one variable at a time, in any order.

7.3 If $f \le g$ are two continuous functions on a box X, prove that $\int_X f \le \int_X g$. Prove that if $f \le g$ are two continuous functions, then $\int_X f = \int_X g$ just when $f = g$.

7.4 If f is a continuous function on a box X and $|f| \le c$ for some constant c, then prove that

$$\left| \int_X f \right| \le c \operatorname{Vol}(X).$$

When does equality occur?

7.5 Prove the fundamental theorem of calculus part (1)

$$\frac{d}{dx} \int_a^x f(u)\, du = f(x)$$

for any continuous function $f \colon [a, b] \to \mathbb{R}$, and part (2)

$$\int_a^x f'(u)\, du = f(x) - f(a)$$

for any continuously differentiable function $f \colon [a, b] \to \mathbb{R}$.

7.6* Suppose that $f(x, y)$ is a function for which

$$\frac{\partial^2 f}{\partial x \partial y} \text{ and } \frac{\partial^2 f}{\partial y \partial x}$$

are defined and continuous. Prove that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Hint: let

$$g(x, y) = \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}.$$

If $g > 0$ at some point, then $g > 0$ on some small box around that point. Why? Integrate g first in x and then in y, and then integrate the other way around.

Notes

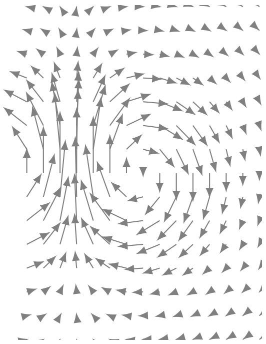
1. The theory of integration we have pursued here is clearly deficient: we can only integrate continuous functions. The Lebesgue theory of integration covers nastier functions, as you will see next year.
2. Improper integrals will play such a small role in these notes that we won't even define them.

Chapter 8

Vector Fields

We define the notion of a vector field and explain how vector fields transform algebraically when we change variables.

Definition

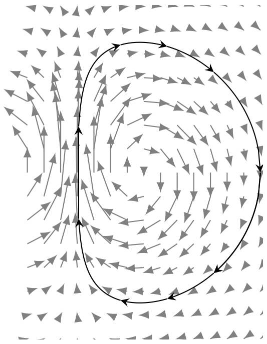


A *vector field* X on an open set $U \subset \mathbb{R}^n$ is a continuously differentiable map

$$X : U \to \mathbb{R}^n.$$

But we think of it not as a map, but as rule which assigns to each point $p \in U$ a vector $X(p)$, which we draw as coming out of the point p.

Picture $X(p)$ as the velocity at the point p of some fluid that fills the region U. If a small particle sat in a fluid flowing this way, how would the particle move? If X is a vector field defined on an open set U, the *flow line* of X through a point $p \in U$ is the curve $x(t)$ so that $x(0) = p$ (start at p), and so that



$$\frac{dx}{dt} = X(x(t))$$

defined for t in a maximal open interval. From your course in ordinary differential equations, you know the existence and uniqueness theorem of ordinary differential equations:

Theorem 8.1 (Picard [1] p. 224)**.** *Every continuously differentiable vector field has unique twice continuously differentiable flow line through each point.*

8.1 Find the flow lines of the vector field on $\mathbb{R}^2$

$$X(x, y) = (y, -x)\,.$$

8.2 Find the flow lines of the vector field on $\mathbb{R}$

$$X(x) = x^k$$

for k a positive integer.

Pushing forward

Recall that a *diffeomorphism* is a continuously differentiable map $F : U \to W$ between open sets $U \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^n$ so that F has a continuously differentiable inverse.

8.3 Prove that the map $F\colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$F(x, y) = (x, y + e^x)$$

is a diffeomorphism.

Given any continuous curve $x(t)$, and diffeomorphism $y = F(x)$, we can of course naturally say that the curve $x(t)$ is taken by F to the curve $y(t) = F(x(t))$.

8.4 What curve in the y variables is the curve

$$(x_1(t), x_2(t)) = (1, t)$$

taken to by the map

$$y = F(x_1, x_2) = (x_1, x_2 + e^{x_1})?$$

Lemma 8.2. *If X is a vector field defined on an open set U, and $F\colon U \to W$ is a diffeomorphism, there is a unique vector field Y on W so that each flow line of X is taken by F to a flow line of Y. This field Y is given by the rule that if $y = f(x)$ then*

$$Y(y) = F'(x)X(x)$$

i.e. applying the matrix $F'(x)$ to the vector $X(x)$.

We denote this vector field F_*X, and call it the *pushforward* of X by F. If X is continuously differentiable and F is twice continuously differentiable, then F_*X is continuously differentiable.

Proof. If there were such a vector field Y, each flow line $x(t)$ would give rise to a flow line $y(t)$ by $y(t) = F(x(t))$, so the velocity of $y(t)$ would be

$$\begin{aligned} y'(t) &= F'(x(t))x'(t) \\ &= F'(x(t))X(x(t)). \end{aligned}$$

But the velocity of a flow line of Y is $y'(t) = Y(y(t))$. Therefore, if Y exists, then $Y(y) = F'(x)X(x)$, i.e.

$$Y(y) = F'\left(F^{-1}(y)\right) X\left(F{-}1(y)\right).$$

So if Y exists, this must be it. If we take this to be Y, then the same steps in reverse tell us that the pushforward of any flow line of X satisfies the equation of a flow line of Y. ☐

8.5 Find F_*e_1 and F_*e_2 where

$$F(x, y) = \left(x, y + x^2\right).$$

Draw pictures.

8.6 Let $F(r,\theta) = (r\cos\theta, r\sin\theta)$, for $r > 0$ and $-\pi < \theta < \pi$. Calculate and draw

$$F_* \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } F_* \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

8.7 Find F_*X where

$$F(x,y) = \left(x^2 - y^2, 2xy\right)$$

for $x > 0$ and

$$X(x,y) = \begin{pmatrix} -y \\ x \end{pmatrix}$$

Draw pictures.

Pushing forward a single vector

We will need to generalize this a little. If F is any continuously differentiable map, and v is any vector, which we imagine lives at a point x, then we can define a corresponding vector $F'(x)v$, which we think of as living at the point $y = F(x)$. We also write this vector as F_*v, if the point x is understood. The difference is that v is just a single vector, not a vector field, and so F_*v is also just a single vector. Moreover, F doesn't have to be a diffeomorphism to define F_*v; F only has to be a continuously differentiable map.

8.8 Let $F \colon \mathbb{R}^2 \to \mathbb{R}^2$ be the map $F(x,y) = \left(xy, y^2\right)$. Find F_*e_1 and F_*e_2 at the point $(1,0)$.

Careful: this process will not necessary generate a vector field in the y-variables. If F is not 1-1, and if X is a vector field, the you might get different values for $F_*X\,(p)$ and $F_*X\,(q)$ even at points p and q with $F(p) = F(q)$.

8.9 Let $F \colon \mathbb{R}^2 \to \mathbb{R}^2$ be the map $F(x,y) = \left(x^2, y^2\right)$. Find F_*e_1 at the points $(x,y) = (1,0)$ and $(x,y) = (-1,0)$.

8.10 Give an example of a continuously differentiable diffeomorphism F and a continuously differentiable vector field X so that F_*X is continuous but not differentiable.

Notes

The perfect introduction to vector fields is Arnol′d's book [1].

Chapter 9

Differential Forms

We define differential forms and their algebraic operations. Differential forms are objects that can be integrated over manifolds, much like integrating functions.

Defining 1-forms

Pick an open set $U \subset \mathbb{R}^n$. A *1-form* on U is a map ω that eats continuous vector fields X on U and splits out continuous functions $\omega(X)$ on U and is "linear over functions", i.e. if f is a continuous function then $\omega\left(fX\right) = f\omega\left(X\right)$, and for any two vector fields $\omega(X+Y) = \omega(X) + \omega(Y)$. For example, we define a 1-form dx_1 by the equation

$$dx_1(X) = X_1,$$

for any vector field

$$X(x) = \begin{pmatrix} X_1(x) \\ X_2(x) \\ \vdots \\ X_n(x) \end{pmatrix}.$$

Similarly we define $dx_2, dx_3, \dots, dx_n$: each component of X is a function.

Lemma 9.1. *Suppose that we write points of $\mathbb{R}^2$ as (x, y). Every 1-form ω on any open set in $\mathbb{R}^2$ can be written uniquely as*

$$\omega = f(x, y)\, dx + g(x, y)\, dy,$$

for some continuous functions $f(x, y)$ and $g(x, y)$.

Proof. Take e_1 and e_2 to be the standard basis vector fields: $e_1(x) = e_1$ for every point x. Then let

$$f(x, y) = \omega\left(e_1\right),$$
$$g(x, y) = \omega\left(e_2\right).$$

Define a differential form

$$\alpha = f(x, y)\, dx + g(x, y)\, dy.$$

We claim $\omega = \alpha$. To check, pick any vector field, say

$$\begin{aligned} X(x,y) &= \begin{pmatrix} X_1(x,y) \\ X_2(x,y) \end{pmatrix}, \\ &= X_1(x,y)e_1 + X_2(x,y)e_2. \end{aligned}$$

Plug in:

$$\begin{aligned} \omega(X) &= X_1(x,y)\omega(e_1) + X_2(x,y)\omega(e_2), \\ &= X_1(x,y)f(x,y) + X_2(x,y)g(x,y) \\ &= \alpha(X). \end{aligned}$$

□

Evaluating a 1-form on a single vector

If v is a vector

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

which sticks out of some point x_0, and ω is a 1-form, say

$$\omega = \sum_i f_i \, dx_i,$$

we let

$$\omega(v) = \sum f_i(x_0) v_i.$$

Another way to say this: if we take any vector field X so that $X(x_0) = v$, then $\omega(v)$ is just $\omega(X)(x_0)$. Clearly this doesn't depend on which vector field X we pick, as long as it takes the right value at x_0, i.e. $X(x_0) = v$.

Higher degree forms

Pick an open set $U \subset \mathbb{R}^n$. A *0-form* on U is just a continuous function $f\colon U \to \mathbb{R}$. For any positive integer k, a *k-form* on U is an operation ω that eats k continuous vector fields, say $X_1, X_2, \dots, X_k$ on U, and spits out a continuous function, which we denote

$$\omega(X_1, X_2, \dots, X_k),$$

so that:

1. The result changes sign if we swap any two of the vector fields.

2. The result is linear, i.e.

   $$\omega\left(Y+Z, X_2, X_3, \dots, X_k\right)=\omega\left(Y, X_2, X_3, \dots, X_k\right)+\omega\left(Z, X_2, X_3, \dots, X_k\right)$$

   for any vector fields Y and Z.
3. The result is "linear over functions", i.e. if f is a continuous function then

   $$\omega\left(f X_1, X_2, \dots, X_k\right)=f \omega\left(X_1, X_2, \dots, X_k\right).$$

A k-form is also called a *differential form* of degree k.

In the plane, define $dx \wedge dy$ to be the 2-form

$$dx \wedge dy\,(X, Y)=dx\,(X)\,dy\,(Y)-dx\,(Y)\,dy\,(X).$$

(Naturally, we want to just use the obvious product formula: the first term. But differential forms must change sign when you swap vector fields, so we need the second term to fix that.) If $\alpha_1, \alpha_2, \dots, \alpha_p$ are any 1-forms, we define a p-form, denoted

$$\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_p$$

by asking that its value on any vector fields $X_1, X_2, \dots, X_p$ be

$$\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_p\left(X_1, X_2, \dots, X_p\right)=\det\begin{pmatrix}\alpha_1\left(X_1\right) & \alpha_1\left(X_2\right) & \dots & \alpha_1\left(X_p\right) \\ \alpha_2\left(X_1\right) & \alpha_2\left(X_2\right) & \dots & \alpha_2\left(X_p\right) \\ \vdots & \vdots & \dots & \vdots \\ \alpha_p\left(X_1\right) & \alpha_p\left(X_2\right) & \dots & \alpha_p\left(X_p\right)\end{pmatrix}.$$

For example, this recovers $dx \wedge dy$ as above.

If we have a list of indices, say $I=(1,3,7,5)$, we write dx_I to mean the form

$$dx_I=dx_1 \wedge dx_3 \wedge dx_7 \wedge dx_5,$$

etc. We can always assume that the numbers in the list increase as we go up the list.

9.1* Prove that these dx_I are linearly independent, in the sense that if there is a linear relation, say

$$0=\sum_I a_I(x) dx_I,$$

where the sum is over all lists of increasing indices, then all of the functions $a_I(x)$ vanish.

Lemma 9.2. *Every differential form ω, say of degree k on an open set in $\mathbb{R}^n$, can be written uniquely as*

$$\omega=\sum_I a_I(x) dx_I,$$

where the sum is over all possible lists I of k indices, say

$$I=\left(i_1, i_2, \dots, i_k\right),$$

a list of increasing integers from 1 to n.

The proof is the same as that of lemma 9.1 on page 57, with more notation.

You won't really need to use this complicated mess, but more generally, if we have any differential forms, say α and β, of degrees a and b, we define a new differential form of degree $a+b$, called the *wedge product* by asking that if $\alpha = \sum_I a_I(x)\, dx^I$ and $\beta = \sum_I b_I(x)\, dx^I$ then

$$\alpha \wedge \beta = \sum_{I,J} a_I(x) b_J(x) dx^{IJ},$$

where IJ means to write out the indices of I and then of J (which might not be strictly increasing indices anymore).

What we need to know is that this wedge product operation satisfies

$$\begin{aligned}
(\alpha_1 + \alpha_2) \wedge \beta &= \alpha_1 \wedge \beta + \alpha_2 \wedge \beta, \\
\alpha \wedge (\beta_1 + \beta_2) &= \alpha \wedge \beta_1 + \alpha \wedge \beta_2, \\
(f\alpha) \wedge \beta &= f\,(\alpha \wedge \beta) \\
&= \alpha \wedge (f\beta)\,, \\
\alpha \wedge \beta &= (-1)^{ab} \beta \wedge \alpha.
\end{aligned}$$

Moreover, the wedge product is associative:

$$\alpha \wedge (\beta \wedge \gamma) = (\alpha \wedge \beta) \wedge \gamma.$$

We won't prove any of these identities; the proofs are just long manipulations of permutations. We also write multiplication by a function, say $f\omega$, as $f \wedge \omega$, treating f as a 0-form.

9.2 Use the identities to prove that $dx \wedge dx = 0$.

9.3 Writing points of $\mathbb{R}^4$ as (x, y, u, v), let $\alpha = dx \wedge dy + du \wedge dv$. Simplify the expression $\alpha \wedge \alpha$.

9.4* Suppose that A is an $n \times n$ matrix, $A = (A_{ij})$. Let $\alpha_i = \sum_j A_{ij} dx_j$, for $i = 1, 2, \dots, n$; these α_i are 1-forms on $\mathbb{R}^n$. Use the identities and induction to prove that

$$\alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_n = \det(A) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

9.5 If $k > n$, prove that the only k-form on any open subset of $\mathbb{R}^n$ is the one that vanishes everywhere: $\omega = 0$. Hint: use lemma 9.2 on the preceding page and problem 9.2.

As for 1-forms, we can evaluate any p-form not only at vector fields but at any p vectors, say $v_1, v_2, \dots, v_p$, sticking out of the same point, say x_0, by

$$\omega\,(v_1, v_2, \dots, v_p) = \omega\,(X_1, X_2, \dots, X_p)|_{x=x_0}$$

for any choice of vector fields $X_1, X_2, \dots, X_p$ so that at x_0

$$X_1\,(x_0) = v_1, X_2\,(x_0) = v_2, \dots, X_p\,(x_0) = v_p.$$

9.6 Suppose that $\alpha_1, \alpha_2, \dots, \alpha_p$ are 1-forms on $\mathbb{R}^n$. Prove that they are linearly independent at every point of $\mathbb{R}^n$ just when $0 \neq \alpha_1 \wedge \alpha_2 \wedge \cdots \wedge \alpha_p$.

Notes

The algebraic manipulation of differential forms seems daunting, when faced with the definition of the wedge product. Actually the algebra is very easy:

1. $dx \wedge dx = 0$, and similarly for any variable in place of x,
2. $dx \wedge dy = -dy \wedge dx$, and similarly for any variables in place of x and y; for example, putting y equal to x gives $dx \wedge dx = 0$,
3. wedge product is associative and distributive over $+$, just like usual multiplication,
4. nothing else: all identities on differential forms are consequences of the above.

Chapter 10

Differentiating Differential Forms

We explain how differential forms transform under maps and how to take the derivative of a differential form.

Pullback

If $F \colon U \to W$ is a continuously differentiable map, $U \subset \mathbb{R}^q$ and $W \subset \mathbb{R}^p$ open sets, and ω is a k-form on W, then the *pullback* of ω by F, denoted $F^*\omega$, is the k-form

$$F^*\omega\left(X_1, X_2, \dots, X_k\right) = \omega\left(F_* X_1, F_* X_2, \dots, F_* X_k\right),$$

by pushing forward the vectors forward at each point. The pullback of any continuously differentiable differential form by a twice continuously differentiable map is a continuously differentiable form. The *pullback* of a function f (in other words, a 0-form) is the function F^*f defined by $F^*f = f \circ F$.

10.1 For the map $(X, Y) = F(x, y) = \left(x + y, y - x^2\right)$, what is the pullback F^*dX? What is F^*dY?

10.2 For the map $F(x, y) = x$, what is the pullback $F^*x^3\,dx$?

10.3 Let U and W be open sets in $\mathbb{R}^n$. Prove that for any diffeomorphism $F \colon U \to W$ and forms α and β on W,

$$F^*\left(\alpha \wedge \beta\right) = \left(F^*\alpha\right) \wedge \left(F^*\beta\right).$$

10.4 Suppose that F and G are two diffeomorphisms. Prove that $\left(F \circ G\right)^*\omega = G^*F^*\omega$ as long as both sides are defined.

Exterior derivative

What is the d in dx? If f is a continuously differentiable function, we define a 1-form df, called the *differential* of f by

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i.$$

We can describe df geometrically: if X is any vector field, with a flow line $x(t)$, then at the point $p = x(0)$,

$$\begin{aligned} df(X) &= \sum_i \frac{\partial f}{\partial x_i} X_i(p), \\ &= \left.\frac{d}{dt}\right|_{t=0} f(x(t)), \end{aligned}$$

by the chain rule. So $df(X)$ at each point is precisely the rate at which the function f changes as we proceed along the flow line of X through that point.

Lemma 10.1. *If $U \subset \mathbb{R}^q$ and $W \subset \mathbb{R}^p$ are open sets and $f \colon W \to \mathbb{R}$ is a continuously differentiable function and $F \colon U \to W$ is a continuously differentiable map, then*

$$F^* df = d\left(F^* f\right).$$

Proof. Pick a vector field X on U, with flow line $x(t)$ through a given point $x_0 \in U$. Let $y_0 = F\left(x_0\right)$ and let $y(t) = F(x(t))$. At the point x_0 we find

$$\begin{aligned} \left(F^* df\right)(X) &= df\left(F_* X\right) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(y(t)) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(F(x(t))) \\ &= d\left(f \circ F\right)(X) \\ &= d\left(F^* f\right)(X). \end{aligned}$$

□

A differential form is *continuously differentiable* if whenever you plug continuously differentiable vector fields into it, you get out a continuously differentiable function. Equivalently, if $\omega = \sum_I a_I(x)\, dx_I$, then ω is continuously differentiable just when all of the $a_I(x)$ functions are continuously differentiable.

If $\omega = \sum_I a_I(x)\, dx_I$, then we denote by $d\omega$ the form

$$d\omega = \sum_I da_I \wedge dx_I = \sum_I \frac{\partial a_I}{\partial x_j} dx_j \wedge dx_I.$$

Theorem 10.2. *The operation d is the unique operation taking continuously differentiable differential forms to differential forms, so that:*

1. *If f is any continuously differentiable function, then df is as defined above.*
2. *For any continuously differentiable differential forms α and β,*

$$d\left(\alpha + \beta\right) = d\alpha + d\beta.$$

3. *The product rule (also called the Leibnitz rule): if* α *is any* a*-form and* β *is any differential form, both continuously differentiable, then*
$$d\left(\alpha \wedge \beta\right) = \left(d\alpha\right) \wedge \beta + (-1)^a \alpha \wedge d\beta.$$
4. *If both of* ω *and* $d\omega$ *are continuously differentiable, then* $d\left(d\omega\right) = 0$. *(This is often expressed by writing* $d^2 = 0$.*)*

Proof. Lets try some examples on the x, y plane to see if there could only be one operation d with those 4 properties. First, we know d on any function f is df. For example, d on the function $f(x, y) = x$ is

$$\begin{aligned} d\left(x\right) &= \frac{\partial x}{\partial x}\,dx + \frac{\partial x}{\partial y}\,dy, \\ &= dx. \end{aligned}$$

On something like $f(x, y)\,dx$ the product rule applies:

$$d\left(f(x, y)\,dx\right) = \left(df\right) \wedge dx + f(x, y)\,d\left(dx\right),$$

since $f(x, y)$ is a 0-form; but $d^2 = 0$ so we get

$$\begin{aligned} &= \left(\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy\right) \wedge dx + 0 \\ &= \frac{\partial f}{\partial y} dy \wedge dx. \end{aligned}$$

The same idea works more generally. Take any differential form ω. By lemma 9.2 on page 59, we can write

$$\omega = \sum_I a_I(x)\,dx_I.$$

We only need to say what d does to each $a_I(x)\,dx_I$ term. But as above,

$$da_I = \sum_j \frac{\partial a_I}{\partial x_j} dx_j,$$

and $d\left(dx_I\right) = 0$. We get the explicit formula

$$d\omega = \sum_j \frac{\partial a_I}{\partial x_i} dx_j \wedge dx_I.$$

So if there is an operation d with the 4 properties listed above, then it must be the operation d that we defined above.

We need to check that this operation d has the 4 required properties. The first two are obvious. For the third property, we can explicitly write out our forms α and β in terms of dx_I and expand, so that it is enough to check $\alpha = dx_I$, and $\beta = dx_J$, for some index lists I and J, for which the result is an easy explicit calculation. Similarly, we can check $d^2 = 0$ explicitly. $\square$

10.5 Use the result of problem 7.6 on page 50 to show that

$$d \circ d\left(f(x) dx_I\right) = 0,$$

for any twice continuously differentiable function f. Explain why this proves that $d^2 = 0$.

10.6 If $\alpha = f(x)\, dx_I$, and $\beta = g(x)\, dx_J$, for some index lists I and J, and I has a elements in it, then check the Leibniz rule, i.e. that

$$d\left(\alpha \wedge \beta\right) = (d\alpha) \wedge \beta + (-1)^a \alpha \wedge d\beta.$$

Hint: you could try induction on a.

Theorem 10.3. *If $F\colon U \to W$ is a continuously differentiable map, $U \subset \mathbb{R}^q$ and $W \subset \mathbb{R}^p$ open sets, then $F^* d\omega = dF^*\omega$, for any continuously differentiable form ω on W.*

Proof. By lemma 10.1 on page 64, this is true if ω is a 0-form. It is also trivially true if ω is $\omega = df$ for some function f, since $d^2 = 0$. If it is true for two differential forms α and β, then it is clearly true for $\alpha \wedge \beta$ by the result of problem 10.3 on page 63. Any differential form is a sum of wedge products of 0-forms and 1-forms of the form df. ☐

Polar coordinates

As an example, consider the map $F\colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$\begin{aligned} (x, y) &= F\left(r, \theta\right), \\ &= \left(r\, \cos\theta, r\, \sin\theta\right). \end{aligned}$$

Then $F^*x = r\, \cos\theta$. We can write this as $x = r\, \cos\theta$. Similarly, $F^*y = r\, \sin\theta$. We can write this as $y = r\, \sin\theta$. The complicated Leibniz rule becomes just the simple product rule on functions:

$$\begin{aligned} dx &= d\left(r \cos\theta\right), \\ &= dr\, \cos\theta + r\, d\left(\cos\theta\right), \\ &= \cos\theta\, dr - r\, \sin\theta\, d\theta. \end{aligned}$$

Similarly,

$$dy = \sin\theta\, dr + r\, \cos\theta\, d\theta.$$

Taking $\wedge$:

$$dx \wedge dy = r\, dr \wedge d\theta.$$

10.7 In polar coordinates, find dr and $d\theta$ in terms of dx and dy.

Notes

1. Differential forms satisfy the algebraic properties that "infinitesimals" should have. We could have used any one of many other definitions.
2. When working In $\mathbb{R}^3$, physicists identify any 1-form

$$\xi = f(x, y, z)\, dx + g(x, y, z)\, dy + h(x, y, z)\, dz$$

with a vector field

$$A = (f(x, y, z), g(x, y, z), h(x, y, z))\,.$$

This identification is invariant under any translation, rotation or reflection. It is not invariant under rescaling, for example. They identify any 2-form

$$\eta = f(x, y, z)\, dy \wedge dz + g(x, y, z)\, dz \wedge dx + h(x, y, z)\, dx \wedge dy,$$

with a vector field

$$F = (f(x, y, z), g(x, y, z), h(x, y, z))\,.$$

This identification is invariant under any translation or rotation, and not under either rescaling or reflection. Finally, they identify any 3-form

$$\omega = f(x, y, z)\, dx \wedge dy \wedge dz$$

with a function $f(x, y, z)$. This identification is again invariant under any translation or rotation, and not under either rescaling or reflection.

These identifications then turn the exterior derivative into an operator on vector fields. If f is a function and we let $\xi = df$, then the corresponding vector field A is denoted

$$A = \nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right),$$

called the *gradient* of f. Similarly if A is any vector field, which we write as

$$A = (A_x, A_y, A_z)$$

with corresponding 1-form ξ, and we let $\eta = d\xi$, then the vector field F corresponding to η is denoted

$$F = \nabla \times A = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}, \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right),$$

called the *curl* of A. Finally, if F is any vector field, written as

$$F = (F_x, F_y, F_z)\,,$$

with corresponding 2-form η, and $\omega = d\eta$, then $\omega = f\, dx \wedge dy \wedge dz$ where f is a function denoted

$$f = \nabla \cdot F,$$

called the *divergence* of F. We will avoid all of this notation, which is clearly special to 3 dimensions and is not invariant under diffeomorphism.

Chapter 11

Integrating Differential Forms

We define the integral of a differential form over an oriented manifold.

Definition

Take an n-form ω on an open set $U \subset \mathbb{R}^n$. We can always write it uniquely as

$$\omega = f(x) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

Take any box $X \subset U$. Define the integral of ω on X to be

$$\int_X \omega = \int_X f.$$

Simple examples of changing variables

We want to compute out the effect of various changes of variable on integrals. For our first example, try $F(x, y) = (y, x)$. Take a 2-form ω on $\mathbb{R}^2$, say $\omega = f(x, y)\, dx \wedge dy$. Suppose that ω has compact support. The pullback is

$$\begin{aligned} F^*\omega &= f(y, x)\, dy \wedge dx, \\ &= -f(y, x)\, dx \wedge dy. \end{aligned}$$

The integral is

$$\int_{\mathbb{R}^2} F^*\omega = -\int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(y, x)\, dy \right) dx,$$

and we just relabel the variables by each other's names:

$$\begin{aligned} &= -\int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(x, y) dx \right) dy, \\ &= -\int_{\mathbb{R}^2} \omega. \end{aligned}$$

In $\mathbb{R}^n$, the same thing happens if we swap any two variables: a sign change. That makes sense because a swap of two variables is orientation reversing.

For our next example, consider $F(x, y) = (-x, y)$. Take a 2-form ω on $\mathbb{R}^2$, say $\omega = f(x, y)\, dx \wedge dy$. Suppose that ω has compact support. The pullback is

$$\begin{aligned} F^*\omega &= f(-x, y)\, d(-x) \wedge dy, \\ &= -f(-x, y)\, dx \wedge dy. \end{aligned}$$

The integral is

$$\int_{\mathbb{R}^2} F^*\omega = -\int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(-x, dy) dx \right) dy$$

use calculus: let $u = -x$

$$= -\int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(u, y)\, du \right) dy,$$

and now we let $x = u$

$$\begin{aligned} &= -\int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(x, y)\, dx \right) dy, \\ &= -\int_{\mathbb{R}^2} \omega. \end{aligned}$$

In $\mathbb{R}^n$, the same thing happens if we change the sign of any variable: a sign change. That makes sense, because a sign change of a variable is orientation reversing.

11.1 Pick a real number $\lambda > 0$. Let $F \colon \mathbb{R} \to \mathbb{R}$ be the map $F(x) = \lambda x$. Consider a compactly supported 1-form $\omega = f(x)\, dx$ on $\mathbb{R}$. Prove that

$$\int_{\mathbb{R}} F^*\omega = \int_{\mathbb{R}} \omega.$$

For our next example, a translation $F(x, y) = (x + x_0, y + y_0)$ in $\mathbb{R}^2$ just acts like changing the variables x and y one at a time by a translation, so we can just look at one variable, and see that

$$\int F^*\omega = \int \omega.$$

For our next example, pick a constant k and consider the linear map

$$F(x, y) = (x, kx + y)$$

given by the matrix

$$\begin{pmatrix} 1 & 0 \\ k & 1 \end{pmatrix}.$$

Take a compactly supported 2-form ω on $\mathbb{R}^2$, say

$$\omega = f(x, y)\, dx \wedge dy.$$

The pullback is

$$\begin{aligned} F^*\omega &= f(x, kx + y)\, dx \wedge d(y + kx), \\ &= f(x, kx + y)\, dx \wedge dy. \end{aligned}$$

The integral is

$$\int_{\mathbb{R}^2} F^*\omega = \int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(x, kx + y)\, dy \right) dx,$$

so for each fixed x, we let $u = kx + y$ and find

$$= \int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(x, u)\, du \right) dx,$$

and now we let $y = u$

$$\begin{aligned} &= \int_{\infty}^{\infty} \left(\int_{\infty}^{\infty} f(x, y)\, dy \right) dx, \\ &= \int_{\mathbb{R}^2} \omega. \end{aligned}$$

The same idea works in any number of variables: if F is any $n \times n$ matrix with 0's everywhere above the diagonal, 1's on the diagonal, and anything below the diagonal, then

$$\int F^*\omega = \int \omega,$$

for any compactly supported n-form ω on $\mathbb{R}^n$. The same idea also works if F is any $n \times n$ matrix with 0's everywhere below the diagonal, and anything above, by the same trick.

An *affine map* is a map of the form $A(x) = c + Tx$ for a linear map T and a constant vector c. The reader can check that A is a diffeomorphism just when $\det T \neq 0$.

Lemma 11.1. *Suppose that $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is an affine diffeomorphism, say $A(x) = c + Tx$, and that ω is a compactly supported n-form on $\mathbb{R}^n$. Then*

$$\int A^*\omega = \pm \int \omega,$$

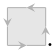
with $\pm = +$ if $\det T > 0$ and $\pm = -$ if $\det T < 0$.

Proof. As we saw above, translations have no effect on integrals over $\mathbb{R}^n$, so assume $c = 0$, and $A(x) = Tx$ is linear. By Gaussian elimination, T can be written as a product of elementary row operation matrices: adding a multiple of a variable to another variable, swapping two variables, and rescaling variables. We know that each of these matrices preserves the integral of ω, except for the appropriate change of sign, so T must as well. ☐

Integrating along the boundary of a box

We can integrate along the boundary of a box, because it is made out of lower dimensional boxes. For example, suppose that $X \subset \mathbb{R}^2$ is a square in the plane. Write

$$\int_{\partial X}$$

to mean that we integrate along the various boundary pieces following the right hand rule: . On any portion of the boundary where x is constant, $dx = 0$, so only the dy part of the integral appears, etc.

11.2 Calculate $\int_{\partial X} x\,dy$, where $X = [0,1] \times [0,1] \subset \mathbb{R}^2$.

If $X = [a,b]$ is a 1-dimensional box, define

$$\int_{\partial X} f = f(b) - f(a),$$

for f any 0-form on X (i.e. any continuous function). In higher dimensions, pick an n-dimensional box $X \subset \mathbb{R}^n$, with $n \geq 1$. Choose a wall of the box. Rotate and translate the box to lie inside the left half space, with our chosen wall along the boundary. (Rotations and translations are orientation preserving, so don't cause any trouble.) Pull back the form by the induced chart along the boundary:

$$(x_2, x_3, \dots, x_n) \mapsto (0, x_2, x_3, \dots, x_n)$$

and integrate. This integral won't change if we rotate the box around, taking that wall to itself, and matching corners to corners. That defines integration along each wall, and we add these up to integrate over ∂X. Clearly the integral over ∂X is invariant under any rotation which takes corners to corners, by definition. This definition is hard to use. Our next lemma will save us from ever having to rotate a box when we carry out an integral, and takes care of all of the minus signs and orientation dilemmas for the rest of the course.

Lemma 11.2. *If X and Y are boxes then*

$$\int_{\partial(X \times Y)} \omega = \int_{\partial X \times Y} \omega + (-1)^{\dim X} \int_{X \times \partial Y} \omega,$$

for any differential form ω defined near $\partial(X \times Y)$.

The notation $\int_{\partial X \times Y}$ means the obvious sum of integrals over the product of each wall of X with Y, with the appropriate orientation on each.

11.3 Why would this help us? It looks like it just restates the problem of orienting a box with the problem of orienting some other boxes.

Proof. For simplicity of notation, take our boxes to be unit boxes: $X = [0,1]^p, Y = [0,1]^q$. Write points of X as x and those of Y as y. and those of $Z = X \times Y$ as $z = (x, y)$. Start with the simplest case: suppose that

$$\begin{aligned} f &= f(x,y), \\ \xi &= dx_2 \wedge \cdots \wedge dx_p, \\ \eta &= dy_1 \wedge \cdots \wedge dy_q, \end{aligned}$$

and let

$$\omega = f\,\xi \wedge \eta.$$

The integral of ξ over any wall of ∂X other than $x_1 = 0$ or $x_1 = 1$ vanishes, because in any other case we are forcing one of the other variables $x_2, x_3, \dots, x_n$ to be constant, say x_i, and then $dx_i = 0$ making $\xi = 0$ along that wall. The same is true of $f\xi \wedge \eta$ integrated along ∂Z: the integral vanishes along any wall of ∂Y, since setting any variable y_j to a constant forces $\eta = 0$. Let

$$\begin{aligned} X_- &= \{\, x \in X \mid x_1 = 0 \,\}, \\ X_+ &= \{\, x \in X \mid x_1 = 1 \,\}, \\ Z_- &= \{\, (x,y) \in Z \mid x_1 = 0 \,\}, \\ Z_+ &= \{\, (x,y) \in Z \mid x_1 = 1 \,\}. \end{aligned}$$

So

$$\int_{\partial X} f\xi = \int_{X_+} f\xi - \int_{X_-} f\xi,$$

and

$$\int_{\partial Z} f\xi \wedge \eta = \int_{Z_+} f\xi \wedge \eta - \int_{Z_-} f\xi \wedge \eta.$$

Along X_+ the chart is

$$(x_2, x_3, \dots, x_p) \mapsto (1, x_2, x_3, \dots, x_p) \in X$$

so that

$$\int_{X_+} f\xi = \int_{[0,1]^{p-1}} f\left(1, x_2, x_3, \dots, x_p, y\right),$$

and by the same reasoning

$$\int_{Z_+} f\xi \wedge \eta = \int_Y \left(\int_{[0,1]^{p-1}} f\left(1, x_2, x_3, \dots, x_p, y\right) \right).$$

The opposite wall has the opposite orientation:

$$\int_{X_-} f\xi = -\int_{[0,1]^{p-1}} f\left(0, x_2, x_3, \dots, x_p, y\right),$$

and

$$\int_{Z_-} f\xi \wedge \eta = -\int_Y \left(\int_{[0,1]^{p-1}} f\left(0, x_2, x_3, \dots, x_p, y\right) \right).$$

Sum these to get

$$\int_{\partial Z} f\xi \wedge \eta = \int_Y \left(\int_{\partial X} f\xi \right) \eta.$$

The next simplest case is if ξ is missing some other dx_2 instead of missing dx_1:

$$\begin{aligned} f &= f(x, y), \\ \xi &= dx_1 \wedge dx_3 \wedge dx_4 \wedge \dots \wedge dx_p, \\ \eta &= dy_1 \wedge \dots \wedge dy_q. \end{aligned}$$

Rotate by a right angle in the x_1, x_2-plane around the center of X and the result holds. The same for dx_3, dx_4, etc.

Next, try

$$\begin{aligned} f &= f(x, y), \\ \xi &= dx_1 \wedge \dots \wedge dx_p, \\ \eta &= dy_2 \wedge \dots \wedge dy_q. \end{aligned}$$

Change variables by the rotation of the x_1, y_1-plane around the center of Z, which replaces y_1 by x_1 and x_1 by $1 - y_1$. Write out the integrals and check the minus signs in $\xi \wedge \eta$. The same trick works for any choice of which dy_i to leave out of η. Every differential form ω of degree $p + q - 1$ (so that we can integrate it over ∂Z) is a sum of differential forms of the various types $f\xi \wedge \eta$ above. ☐

Stokes's theorem in a box

Theorem 11.3 (Stokes's theorem in a box)**.** *Suppose that $X \subset \mathbb{R}^n$ is a box and ω is a continuously differentiable $(n-1)$-form defined on X. Then*

$$\int_X d\omega = \int_{\partial X} \omega.$$

Proof. Let's do the proof for the box

$$X = [0, 1] \times [0, 1] \subset \mathbb{R}^2$$

in the plane. (The general proof is just more notation.) We can write any 1-form in the plane as

$$\omega = f(x, y)\, dx + g(x, y)\, dy.$$

It is enough to separately prove the result for each of $f(x,y)\,dx$ and $g(x,y)\,dy$. We will do the first; the second follows by the same method, just switching the roles of x and y. So assume that

$$\omega = f(x,y)\,dx$$

so that

$$d\omega = -\frac{\partial f}{\partial y} dx \wedge dy$$

Compute

$$\begin{aligned} \int_X d\omega &= \int_X \left(-\frac{\partial f}{\partial y}\right) dx \wedge dy, \\ &= -\int_0^1 \int_0^1 \frac{\partial f}{\partial y}\, dx\, dy, \end{aligned}$$

and by the fundamental theorem of calculus this is

$$\begin{aligned} &= -\int_0^1 \left(f(x,1) - f(x,0)\right)\, dx, \\ &= \int_0^1 f(x,0)\, dx - \int_0^1 f(x,1)\, dx, \end{aligned}$$

and we leave the reader to check that this is precisely

$$= \int_{\partial X} \omega.$$

□

Review problems

11.4 A differential form ω is *exact* if $\omega = d\xi$ for some continuously differentiable differential form ξ. Find an example of a differential form which is not exact.

11.5 A continuously differentiable differential form ω is *closed* if $d\omega = 0$. For example, on $\mathbb{R}^2 - \{0\}$, the 1-form

$$\omega = \frac{-y\,dx + x\,dy}{x^2 + y^2}$$

is closed, but turns out not to be exact, as we will see. Prove that continuously differentiable exact forms are closed.

11.6* If ω is an n-form in $\mathbb{R}^n$, vanishing for $x_1 \le 0$, then prove that ω is exact, and that $\omega = d\xi$ for some continuously differentiable $(n-1)$-form vanishing for $x_1 \le 0$.

11.7 Suppose that $U \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^n$ are open sets. Suppose that $F\colon U \to W$ and $G\colon U \to W$ are continuously differentiable maps. Take a box $X \subset U$. Suppose that near ∂X, $F = G$. Use the result of problem 11.6 on the preceding page to prove that

$$\int_X F^* \omega = \int_X G^* \omega$$

for any n-form defined on W.

11.8 In one dimension, i.e. $n = 1$, solve the last problem more directly.

11.9* Suppose that X is a box and $F\colon X \to X$ is a continuously differentiable map and $F(x) = x$ for all $x \in \partial X$. Use the result of problem 11.7 to prove that F is onto. Hint: take $G(x) = x$, and take lots of "bump functions" f, letting

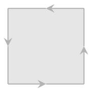
$$\omega = f(x) dx_1 \wedge dx_2 \wedge \dots \wedge dx_n.$$

11.10 Give a short proof of lemma 11.2 on page 72 for the simplest case: $X = Y = [0, 1]$.

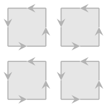
11.11 Do you see any (vague) similarity between the boundary operation ∂ and the exterior derivative operation d?

Notes

If we cut up an oriented box



into a grid, the grid boxes orient like:



(We insert a little space between the grid boxes to make room to draw the arrows, but there really isn't any such space.) When two grid boxes meet, they meet along a pair of sides that run in opposite directions. The integrals over these sides cancel each other out. Therefore the integral over the boundary of the original box is the sum of the integrals over the grid boxes.

Chapter 12

Stokes's Theorem

We prove Stokes's theorem for manifolds with boundary and corners, and prove that integrals of differential forms are invariant under diffeomorphism.

Change of variable in integrals

How can we make global problems into local problems? Suppose that we have a differential form ω defined near a compact set $K \subset \mathbb{R}^n$. Take any open cover of K; imagine that the cover consists of very small balls for example. Then pick a partition of unity $f_1, f_2, \dots, f_k$ subordinate to that open cover. At each point $x \in K$, we have

$$1 = f_1(x) + f_2(x) + \dots + f_k(x).$$

Therefore on K,

$$\omega = f_1\omega + f_2\omega + \dots + f_k\omega.$$

So ω is a sum of forms, each one of which is supported on one of the open sets in our cover (i.e. in one very small ball). To prove various theorems about ω, it will often be enough to prove those theorems about each of $f_1\omega, f_2\omega, \dots, f_k\omega$. We can add up local solutions to various problems to get a global solution.

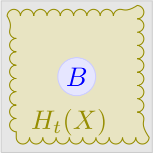
Lemma 12.1. *Suppose that $F \colon U \to W$ is a positively oriented diffeomorphism of open sets of $\mathbb{R}^n$ and ω is a compactly supported n-form on W. Then*

$$\int F^*\omega = \int \omega.$$

Proof. Covering the set where $\omega \neq 0$ in small open sets, and using a partition of unity, it is clear that we only need to prove our result locally, i.e. for n-forms ω vanishing outside some arbitrarily small ball, say around some point $y_0 = F(x_0)$. By composing F with affine maps as needed, we can arrange that $x_0 = y_0 = 0$ and $F'(0) = I$. So $F(x) = x + Q(x)$, say, where $Q(x) = o(x)$. In particular, F only slightly modifies the shape of a small enough box X around the origin, as we saw in problem 3.12 on page 27. Let $H(t,x) = x + tQ(x)$, for any $x \in X$ and any $t \in \mathbb{R}$. For each t, let $H_t(x) = H(t,x)$. If we perhaps replace X by a smaller box, say with sides of length δ, we can arrange that $\|Q(x)\| \le \varepsilon \|x\|$, for any $x \in X$, for any ε we like. So for any t with $0 \le t \le 1$, H_t moves the points of X only very slightly, by at most $\varepsilon\delta$. Once again using a partition of unity, we can arrange that ω vanishes everywhere outside a very small ball $B \subset X$.

We can assume that B is as small as we like, so that all of the maps H_t, for $0 \le t \le 1$, send ∂X to points outside B:

$H_0(X) = X$



Since ω is an n-form on $\mathbb{R}^n$, clearly $d\omega = 0$, so

$$\begin{aligned} 0 &= \int_{[0,1]\times X} H^* d\omega, \\ &= \int_{[0,1]\times X} dH^*\omega, \end{aligned}$$

to which we apply Stokes's theorem

$$= \int_{\partial([0,1]\times X)} H^*\omega,$$

to which we apply lemma 11.2

$$= \int_{\{1\}\times X} H^*\omega - \int_{\{0\}\times X} H^*\omega - \int_{[0,1]\times \partial X} H^*\omega,$$

If $x \in \partial X$ and $0 \le t \le 1$, then $H(t,x)$ lies outside B. Outside B, $\omega = 0$, so $H^*\omega = 0$ on $[0,1] \times \partial X$, killing that last term:

$$= \int_{\{1\}\times X} H^*\omega - \int_{\{0\}\times X} H^*\omega,$$

but $H_1 = F$ and $H_0(x) = x$ so

$$\begin{aligned} &= \int_X F^*\omega - \int_X \omega, \\ &= \int F^*\omega - \int \omega. \end{aligned}$$

□

12.1 To go back to integrating functions, use the previous result to prove that

$$\int_X f(F(x)) \left|\det F'(x)\right| = \int_{F(X)} f(y),$$

for $F \colon X \to F(X) \subset \mathbb{R}^n$ a diffeomorphism on a box and $f \colon F(X) \to \mathbb{R}$ a continuous function.

Lemma 12.2. *Suppose that $F \colon U \to W$ is a positively oriented diffeomorphism of open sets of the left half space, or of a box, and ω is a compactly supported differential form on W. Then*

$$\int F^* \omega = \int \omega.$$

Proof. Consider open sets U and W in an n-dimensional box and an n-form ω on W. If $\omega = 0$ near all of the walls, then the proof is the same as above. So imagine that $\omega \neq 0$ at some point on some wall. Pick a continuous function f, vanishing in a thin layer near the walls, say a layer of some thickness everywhere at most δ_1, then increasing up to 1 in another thin layer, say of thickness everywhere at most δ_2, and then staying 1 everywhere further away from the walls. By the above reasoning,

$$\int F^* f\omega = \int f\omega.$$

But clearly $\int f\omega$ is as close as we like to $\int \omega$, as $\delta_1, \delta_2 \to 0$, and $\int F^* f\omega$ is as close to $\int F^* \omega$, since the layers get mapped to layers of thickness at most $\delta_1 D$ and $\delta_2 D$, where $D = \max_x \|F'(x)\|$, the maximum taken over the points x lying in those layers and at which $\omega \neq 0$. $\square$

Suppose that $M \subset \mathbb{R}^n$ is a p-dimensional oriented manifold, perhaps with boundary or with corners, and that ω is a p-form defined near M, vanishing at every point of M except on some compact set $K \subset M$. We want to define $\int_M \omega$. Consider the open cover of K consisting of all open sets $U_\alpha \subset \mathbb{R}^n$ so that $U_\alpha \cap M$ lies inside the image of a single chart. Pick a partition of unity $f_1, f_2, \dots, f_N$ subordinate to that open cover. So each f_i vanishes at every point of M outside the image of a single chart, say F_i. Define

$$\int_M \omega = \sum_i \int F_i^* \left(f_i \omega\right).$$

Lemma 12.3. *The integral $\int_M \omega$ defined above does not depend on the choice of partition of unity or charts.*

Proof. Suppose that you pick a partition of unity $f_1, f_2, \dots, f_M$ and some charts $F_1, F_2, \dots, F_N$ so that $f_i \omega$ vanishes outside the image of F_i. I pick another partition of unity, say $g_1, g_2, \dots, g_N$, and some charts $G_1, G_2, \dots, G_N$ so that $g_j \omega$ vanishes outside the image of G_j. For any i and j, consider the 1-form

$$f_i g_j \omega.$$

By lemma 12.1 on page 77,

$$\int F_i^* \left(f_i g_j \omega\right) = \int G_j^* \left(f_i g_j \omega\right).$$

But then

$$\sum_{ij} \int F_i^* \left(f_i g_j \omega\right) = \sum_{ij} \int G_j^* \left(f_i g_j \omega\right).$$

Clearly the left hand side is

$$\sum_i \int F_i^* \left(\sum_j f_i g_j \omega\right) = \sum_i \int F_i^* \left(f_i \omega\right),$$

while the right hand side is

$$\sum_j \int G_j^* \left(\sum_i f_i g_j \omega\right) = \sum_j \int G_j^* \left(g_j \omega\right).$$

□

Lemma 12.4. *If F is a diffeomorphism defined near an oriented manifold M (perhaps with boundary and corners), then $\int_M F^*\omega = \int_{F(M)} \omega$ for any differential form ω defined near $F(M)$ whose support intersects $F(M)$ in a compact set.*

Proof. Take charts F_i as above for M, and $F \circ F_i$ as charts for $F(M)$. □

12.2 Use a change of variables to compute the area inside the ellipse

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

using the well-known area inside a circle.

Stokes's theorem

Theorem 12.5 (Stokes's theorem)**.** *Suppose that $M \subset \mathbb{R}^n$ is a manifold, perhaps with boundary or corners, and ω is a continuously differentiable differential form defined on an open set containing M. Suppose that ω vanishes near every point of M except on some compact set. Then*

$$\int_M d\omega = \int_{\partial M} \omega.$$

Proof. It works in a box, and any integral is a sum of integrals over boxes. □

12.3 We can use integrals around curves to compute areas. Prove that for M any open set in $\mathbb{R}^2$ with boundary and corners, the area of M is

$$\int_{\partial M} x \, dy.$$

Use this to compute the area of a disk of radius R.

12.4 Use Stokes's theorem to find the area inside the curve

$$x^{2/5} + y^{2/5} = 1.$$

12.5 Take a positive continuously differentiable function $y = f(x)$ of one variable, and two real numbers $a < b$. Let M be the manifold with boundary and corners consisting of all points (x, y) with $a \le x \le b$ and $0 \le y \le f(x)$. Draw a picture. Use Stokes's theorem applied to $\omega = y\,dx$ to recover the usual procedure for calculating area under the graph of a function from one variable calculus. Be careful with minus signs.

The Cauchy integral theorem

This section can be omitted without loss of continuity.

If M is an n-dimensional manifold with boundary and corners in $\mathbb{R}^n$, then the open set $U = M - \partial M$ is called a *domain*. We write ∂U to mean ∂M and call ∂U the *boundary* of U.

Let $\mathbb{C}$ be the set of complex numbers, with $i = \sqrt{-1}$. Each complex number z splits uniquely into $z = x + iy$, with x and y real, and this identifies $\mathbb{C} = \mathbb{R}^2$. Write $\bar{z}$ to mean $x - iy$, and write dz to mean $dx + i\,dy$, etc.

Every complex-valued 1-form can be written in terms of dx and dy, with complex number coefficients. Solve the equations

$$\begin{aligned} dz &= dx + i\,dy, \\ d\bar{z} &= dx - i\,dy, \end{aligned}$$

to find that

$$\begin{aligned} dx &= \frac{1}{2}\,dz + \frac{1}{2}\,d\bar{z}, \\ dy &= -\frac{i}{2}\,dz + \frac{i}{2}\,d\bar{z}. \end{aligned}$$

Therefore every complex valued differential form on any open subset of $\mathbb{C}$ can be expressed in terms of dz and $d\bar{z}$. Each complex valued 1-form looks like

$$\omega = f(x, y)\,dz + g(x, y)\,d\bar{z},$$

for some continuous functions $f(x, y)$ and $g(x, y)$. We take exterior derivative of a complex valued form by

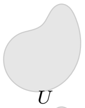
$$d\,(\xi + i\,\eta) = d\xi + i\,d\eta,$$

for the real and imaginary parts ξ and η. A continuously differentiable complex valued function f is *holomorphic* if $df \wedge dz = 0$.

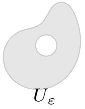
Theorem 12.6. *Suppose that $U \subset \mathbb{C}$ is a bounded domain with continuously differentiable boundary and corners. Then*

$$f(z) = \frac{1}{2\pi i} \int_{\partial U} \frac{f(w)}{w-z}\, dw,$$

for every point $z \in U$ and every holomorphic function f on $\bar{U}$. In particular, the holomorphic function f is completely determined by its values on ∂U.



U



U_ε

Proof. Let

$$\omega = \frac{1}{2\pi i} \frac{f(w)}{w-z}\, dw.$$

Calculate $d\omega = 0$. Pick a small value of $\varepsilon > 0$ and let $U_\varepsilon = U - B_z(\varepsilon)$.

By Stokes's theorem,

$$\int_{\partial U_\varepsilon} \omega = 0.$$

If ε is small enough that $B_0(\varepsilon)$ does not intersect ∂U, then

$$\partial U_\varepsilon = \partial U \cup \partial B_0(\varepsilon),$$

with opposite orientations on the two boundary components, so

$$\begin{aligned} \int_{\partial U} \omega &= \int_{\partial B_0(\varepsilon)} \omega, \\ &= \frac{1}{2\pi i} \int_{\partial B_0(\varepsilon)} \frac{f(w)}{w-z}\, dz, \end{aligned}$$

to which we apply the change of variable $w = z + \varepsilon e^{i\theta}$ to get

$$\begin{aligned} &= \frac{1}{2\pi i} \int_0^{2\pi} \frac{f\left(z+\varepsilon e^{i\theta}\right)}{\varepsilon e^{i\theta}}\, i\varepsilon e^{i\theta}\, d\theta, \\ &= \frac{1}{2\pi} \int_0^{2\pi} f\left(z+\varepsilon e^{i\theta}\right)\, d\theta, \end{aligned}$$

just giving the average value of f over the circle of radius ε around z, which by continuity approaches $f(z)$ as $\varepsilon \to 0$. ☐

Review problems

12.6 For those who like long calculations: compute

$$\int_X \left(1 - \sqrt{x^2+y^2+z^2}\right) dx\, dy\, dz$$

over the hemisphere

$$X = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; z \ge 0,\ x^2+y^2+z^2 \le 1 \right\}.$$

Hint: try *spherical coordinates*

$$\begin{aligned} x &= r\,\sin\alpha\cos\beta, \\ y &= r\,\sin\alpha\sin\beta, \\ z &= r\,\cos\alpha. \end{aligned}$$

12.7 Let $X \subset \mathbb{R}^3$ be the set of points (x, y, z) bounded by the cylinders $x^2 + y^2 \geq 4$, $x^2 + y^2 \leq 9$, and with $0 \leq z \leq 1$ and with $0 \leq y \leq x$. Draw X. Using cylindrical coordinates

$$\begin{aligned} x &= r\,\cos\theta, \\ y &= r\,\sin\theta, \\ z &= z, \end{aligned}$$

i.e. the map $F(r, \theta, z) = (r\,\cos\theta, r\,\sin\theta, z)$, compute

$$\int_X z\left(x^2 + y^2\right)\,dx\,dy\,dz.$$

12.8 Suppose that S is an n-dimensional manifold with boundary and corners, ω is a continuously differentiable $(n-1)$-form defined near S and f is a continuously differentiable function defined near S. Suppose that $f\omega$ vanishes everywhere except in a compact set. Prove the equation of integration by parts:

$$\int_S f\,d\omega = \int_{\partial S} f\omega - \int_S df \wedge \omega.$$

Explain how this generalizes integration by parts from one variable calculus.

Notes

1. Imagine a simple but geometric definition of integration of differential forms on manifolds: cut the manifold into small pieces, each piece being the image of a chart. Then pull back the form by each chart, and just integrate in Euclidean space. This naive definition is not so easy. You need to show that the result doesn't depend on how you cut up the manifold. There are many ways to cut up a manifold. Partitions of unity are a trick to avoid cutting manifolds. Instead we work with functions (or forms) defined on a single chart first. Then we add up such functions or forms. Existence of a partition of unity ensures that we can build up any function or form out of these local contributions.
2. The term *domain* is used in analysis to mean any open set which is equal to the interior of its closure. For example, a punctured ball is not a domain. We will only want to consider domains with continuously differentiable boundary with corners, an awkward phrase, so we have adopted a more narrow definition of the term.

Chapter 13

The Brouwer Fixed Point Theorem

We prove a fundamental result in topology: any map from a closed ball to itself has a fixed point.

The angle that a curve winds around the origin

Recall polar coordinates: $F\colon \mathbb{R}^2 \to \mathbb{R}^2$

$$(x, y) = F(r, \theta) = (r\,\cos\theta, r\,\sin\theta)\,.$$

This F is not a diffeomorphism for 2 reasons:

1. $F(r,\theta) = F(r, \theta + 2\pi)$, so different inputs with same output, and
2. check that $F'(0,\theta) = 0$.

Recall

$$\begin{aligned} F^* dx &= \cos\theta\,dr - r\,\sin\theta\,d\theta, \\ F^* dy &= \sin\theta\,dr + r\,\cos\theta\,d\theta. \end{aligned}$$

The function θ cannot be expressed in terms of x and y, because as we said $F(r, \theta + 2\pi) = F(r, \theta)$. Of course $r = \sqrt{x^2 + y^2}$. But strangely $d\theta$ is expressible in terms of x and y:

$$d\theta = F^* \frac{-y\,dx + x\,dy}{x^2 + y^2}.$$

Therefore we always write $d\theta$ to mean $\frac{-y\,dx + x\,dy}{x^2+y^2}$. If $(x(t), y(t))$ is a curve winding around the origin, then the total angle it winds is

$$\int d\theta.$$

13.1 Calculate the integral of $d\theta$ around a circle centered at the origin. Why is $d\theta$ not exact (i.e. not really d of something) in the x, y-plane minus the origin?

To compute integrals of areas of regions X in the plane, we need to integrate

$$\int_X dx \wedge dy.$$

If we can write X in terms of polar coordinates, it might be easier.

13.2 How to compute areas: calculate that

$$F^* dx \wedge dy = r\, dr \wedge d\theta.$$

13.3 Find the area of the disk of radius R in the plane around the origin, using polar coordinates.

13.4 Calculate the area of the disk by using Stokes's theorem applied to the 1-form $\xi = y\, dx$ and computing $\int \xi$ around the boundary of the disk.

Hook

If you plug just one vector field into a 2-form, you are left with a 1-form:

$$dx \wedge dy\, (e_1, \ \ ) = dy.$$

If ω is a p-form and X is a vector field, we denote by $X \lrcorner\, \omega$ the $(p-1)$-form

$$\omega\, (X, \ \ , \dots, \ \ )\,.$$

In other words, $X \lrcorner\, \omega$ is the $(p-1)$-form that eats any vector fields $Y_1, Y_2, \dots, Y_{p-1}$ and splits out

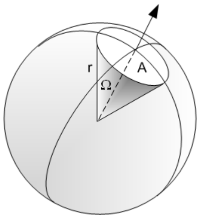
$$X \lrcorner\, \omega\, (Y_1, Y_2, \dots, Y_{p-1}) = \omega\, (X, Y_1, Y_2, \dots, Y_{p-1})\,;$$

just stick an X in the first slot, and you have a $(p-1)$-form.

The *solid angle form* Ω on $\mathbb{R}^n$ is the $(n-1)$-form which, at any point $x \in \mathbb{R}^n - \{0\}$, is

$$\Omega = \frac{1}{\|x\|^n} x \lrcorner\, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n.$$

We will see that its integral over a piece of the unit sphere is the "solid angle" of that piece:



13.5 Write out Ω in $\mathbb{R}^3$ in the usual x, y, z coordinates.

13.6 Using the result of the previous problem, write out Ω in $\mathbb{R}^3$ in spherical coordinates:

$$\begin{aligned} x &= r\, \sin\alpha \cos\beta, \\ y &= r\, \sin\alpha \sin\beta, \\ z &= r\, \cos\alpha. \end{aligned}$$

13.7 Prove that $d\mathit{\Omega} = 0$.

13.8 For $x \in \mathbb{R}^n$, let $r = \|x\|$. Prove that

$$d\left(r^2\right) \wedge \mathit{\Omega} = 2r^{3-n}\, dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

Conclude that in "polar coordinates" in $\mathbb{R}^n$

$$dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n = r^{n-1} dr \wedge \mathit{\Omega}.$$

13.9 Prove that the number

$$\int_S \mathit{\Omega}$$

is the same number for any sphere S which surrounds the origin, and that it is 0 for any sphere which is the boundary of a ball not containing the origin.

Take an open set $U \subset S$ of the unit sphere S. For any two radii $R_1 < R_2$, let U_{R_1,R_2} be the set of points x so that $R_1 < \|x\| < R_2$ and

$$\frac{x}{\|x\|} \in U.$$

Clearly U_{R_1,R_2} is an n-dimensional manifold.

$$\begin{aligned}
\operatorname{Vol}\left(U_{R_1,R_2}\right) &= \int_{U_{R_1,R_2}} r^{n-1} dr \wedge \mathit{\Omega}, \\
&= \int_{R_1}^{R_2} r^{n-1} dr \int_U \mathit{\Omega}, \\
&= \left.\frac{r^n}{n}\right|_{R_1}^{R_2} \int_U \mathit{\Omega}, \\
&= \frac{R_2^n - R_1^n}{n} \int_U \mathit{\Omega}.
\end{aligned}$$

In particular, $\int_U \mathit{\Omega} > 0$. We call $\int_U \mathit{\Omega}$ the *solid angle* of a subset U of the sphere.

13.10 Sending $R_1 \to 0$ and setting $R_2 = 1$, draw a picture to explain where the solid angle form gets its name.

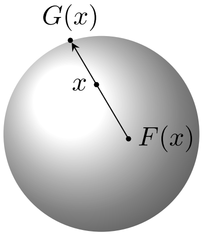
13.11 Check that the solid angle of a piece of a circle around the origin is the usual notion of angle.

13.12 What is the solid angle of the entire unit sphere in $\mathbb{R}^3$?

The Brouwer fixed point theorem

Theorem 13.1. *Let $\bar{B}$ be a closed ball in $\mathbb{R}^n$. Suppose that $F\colon \bar{B} \to \bar{B}$ is a continuously differentiable map. Then F has a fixed point.*

Proof. Suppose not. Assume that $\bar{B} = \bar{B}_1(0)$. Assume that F has no fixed point. Let $S = \partial B$. Take any point $x \in \bar{B}$. Since $F(x) \neq x$, the two points x and $F(x)$ lie on a unique line. Follow that line starting from $F(x)$ in the direction of x until you get to a point of S. Call that point $G(x)$.



This defines a map

$$G\colon \bar{B} \to S,$$

so that $G(x) = x$ for $x \in S$. Also note that $\|G(x)\| = 1$ for any $x \in \bar{B}$. We let the reader prove that G is continuously differentiable on $\bar{B}$. Compute solid angles:

$$0 < \int_S \mathit{\Omega},$$

but $G(x) = x$ on S so

$$= \int_S G^* \mathit{\Omega},$$

to which we apply Stokes's theorem

$$\begin{aligned} &= \int_{\bar{B}} dG^* \mathit{\Omega}, \\ &= \int_{\bar{B}} G^* d\mathit{\Omega}, \\ &= 0 \end{aligned}$$

because $d\mathit{\Omega} = 0$. $\square$

13.13* Prove that the map G described in the proof of the Brouwer fixed point theorem is well defined and continuously differentiable.

13.14 In the proof of the Brouwer fixed point theorem: the form $\varOmega$ is not defined at 0, so is only defined on $\mathbb{R}^n - \{0\}$. So why is $G^*\varOmega$ defined and continuously differentiable on all of $\bar{B}$?

13.15 The Brouwer fixed point theorem holds true for merely continuous maps, not necessarily continuously differentiable. Use that fact (which we haven't proven) in this problem.

1. Take two identical pieces of graph paper of the same size and shape. Pick up one piece of paper and crumple it. Put it on top of the other piece of paper. Prove that there must be a point of the crumpled piece of paper which is now lying exactly on top of the corresponding point of the lower sheet. (Assume that crumpling is continuous.)
2. Take a still cup of coffee and stir it. Wait until the coffee is still again. Prove that some point inside the fluid of the coffee is now located exactly where it was before you stirred the coffee. (Assume that stirring is continuous.)

13.16 Suppose that X is a continuously differentiable vector field on the unit closed ball $\bar{B} \subset \mathbb{R}^n$. Suppose that $X(x) = x$ for all $x \in \partial\bar{B}$; in other words X points "straight out of the sphere" at all points of the unit sphere. Prove that X has a zero somewhere inside the ball $\bar{B}$. Hint: use the map

$$F(x) = \frac{X(x)}{\|X(x)\|},$$

integration of the solid angle form $\varOmega$, Stokes's theorem, and the fact that $d\varOmega = 0$.

Notes

The Brouwer fixed point theorem is also true for merely continuous maps, but needs a different proof. The Brouwer fixed point theorem differs from the contraction mapping theorem because the map in Brouwer's theorem might stretch some places and squish others, just as long as it fits the closed ball inside itself. A contraction map has to squish everywhere, stretch nowhere.

Chapter 14

Manifolds from Inside

We start to look at a manifold as if we live inside it.

Tangent vectors

In a picture, we think of a tangent vector to a manifold as some vector just glancing off that manifold: A vector $v \in \mathbb{R}^n$ is *tangent* to a manifold M at a point $x_0 \in M$ if there is a continuously differentiable curve $x(t) \in M$, say for t in some neighborhood of $0 \in \mathbb{R}$, with $x(0) = x_0$ and $x'(0) = v$. The set of all tangent vectors at a point $x_0 \in M$ is called the *tangent space* to M at x_0, denoted $T_{x_0} M$.

Lemma 14.1. *The tangent space to a p-dimensional manifold $M \subset \mathbb{R}^n$ at a point x_0 is a p-dimensional linear subspace of $\mathbb{R}^n$. Moreover, for any chart h, if $h(u_0) = x_0$, then*

$$T_{x_0} M = \operatorname{image} h'(u_0).$$

Proof. Take any curve $x(t) \in M$ with $x(0) = x_0$. Let $u(t) = h^{-1}(x(t))$, so

$$x'(0) = h'(u_0)\, u'(0) \in \operatorname{image} h'(u_0).$$

Conversely, if $v \in \operatorname{image} h'(u_0)$, say $v = h'(u_0)\, w$, then we can take the curve $u(t) = u_0 + tw$, and let $x(t) = h(u(t))$. ☐

14.1 Find the tangent space at a point $(x, 0, z)$ to the hyperboloid $x^2 + y^2 = z^2 + 1$.

14.2 Prove that

$$T_{(m,n)}(M \times N) = T_m M \oplus T_n N$$

for any points $m \in M$ and $n \in N$ for any manifolds M and N.

14.3 Find the tangent space at a point (x, y, z) to the graph of a function $z = f(x, y)$.

Derivatives

To check if a map $f\colon M \to \mathbb{R}^p$ on a manifold M is continuously differentiable, it is enough to check, for some chart h for M, if $f \circ h$ is continuously differentiable, and check any one chart near each point of M.

For example, on the circle in the plane, the functions x and y are continuously differentiable, because for any constant α_0, the chart $h\colon t \mapsto (\cos t, \sin t)$ defined for $\alpha_0 < t < \alpha_0 + 2\pi$ has

$$x \circ h(t) = \cos t, y \circ h(t) = \sin t,$$

continuously differentiable.

A map $F\colon P \to Q$ between manifolds $P \subset \mathbb{R}^p$ and $Q \subset \mathbb{R}^q$ is *continuously differentiable* if the composition of maps $P \to Q \to \mathbb{R}^q$ is.

The *derivative* of a continuously differentiable map $F\colon P \to Q$ at a point $x_0 \in P$ is the linear map

$$F'\left(x_0\right) : T_{x_0} P \to T_{y_0} Q,$$

where $y_0 = F\left(x_0\right)$, defined by $F'\left(x_0\right) x'(t) = y'(t)$ whenever $x(t)$ is a continuously differentiable curve in P and $y(t) = F(x(t))$.

Lemma 14.2. *Any continuously differentiable map $f\colon P \to Q$ between manifolds has a unique derivative at each point of P.*

Proof. Since the possible values of $x'(t)$ span the tangent space of P, there is only at most one linear map $f'(x)$ that can satisfy the equation $y'(t) = f'(x(t))x'(t)$ for every curve $x(t)$. So $f'(x)$ is uniquely determined by this equation, if it exists.

Pick a chart h for P. For each x in the image of h, let

$$f'(x) = (f \circ h)' \left(h^{-1}(x)\right) h' \left(h^{-1}(x)\right)^{-1}.$$

We need to prove that this choice of $f'(x)$ is a derivative. Pick a curve $x(t) \in P$ and let $y(t) = f(x(t))$. We can assume that the curve stays inside the image of h; $u(t) = H^{-1}(x(t))$ is continuously differentiable and $u'(t) = h'(u(t))^{-1} x'(t)$. Therefore $x'(t) = h'(u(t))u'(t)$, while $y(t) = f(h(u(t)))$, so

$$\begin{aligned} y'(t) &= (f \circ h)' (u(t)) \, u'(t), \\ &= (f \circ h)' \left(h^{-1}(x(t))\right) h' \left(h^{-1}(x)\right)^{-1} x'(t), \\ &= f'(x(t))x'(t). \end{aligned}$$

□

A *tangent vector field* (also often just called a *vector field*) on a manifold M is a continuous map $X\colon M \to \mathbb{R}^n$ so that $X(x) \in T_x M$ for every point $x \in M$.

The derivative of a continuously differentiable map is continuous in the sense that, if X is any continuous tangent vector field, then $x \mapsto f'(x)X(x)$ is a continuous map.

Differential forms on manifolds

A *differential form* ω of degree p on a manifold $M \subset \mathbb{R}^n$ is an operation that eats p tangent vector fields $X_1, X_2, \dots, X_p$ on M and spits out a continuous function

$$\omega\left(X_1, X_2, \dots, X_p\right),$$

so that:

1. The result changes sign if we swap any two of the vector fields.
2. The result is “linear over functions”, i.e. if f is a continuously differentiable function, then

$$\omega\left(f X_1, X_2, \dots, X_k\right) = f \omega\left(X_1, X_2, \dots, X_k\right).$$

Theorem 14.3 (Stokes's theorem)**.** *If ω is a continuously differentiable $(n-1)$-form on an n-dimensional manifold M, perhaps with boundary and corners, then*

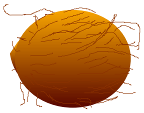
$$\int_{\partial M} \omega = \int_M d\omega.$$

The proof is clear: it works in each chart as before, and you just add up using a partition of unity.

The coconut theorem

Theorem 14.4. *A sphere admits a nowhere vanishing tangent vector field just when the sphere has odd dimension.*

For example, any sphere in $\mathbb{R}^3$ has dimension 2, so every tangent vector field to the sphere must have a zero. In other words, if a sphere has hair growing out of it (like a coconut), then you can't comb all of the hair so that it all lies tangent to the sphere.



Proof. Suppose that the sphere is $S^n \subset \mathbb{R}^{n+1}$. Suppose that there is a tangent vector field X on S nowhere vanishing. Check that $X(x)$ is perpendicular to x, at every point $x \in S^n$. Replace X by $X/\left\|X\right\|$ to arrange that $\|X\| = 1$, a unit tangent vector field. Let $M = [0,1] \times S^n$, a manifold with boundary inside $\mathbb{R}^{n+2}$. Define a map

$$F\colon (t,x) \in M \mapsto \cos(\pi t)x + \sin(\pi t)X(x) \in \mathbb{R}^{n+1}.$$

Check that $\|F(x,t)\| = 1$ for every point $x \in S^n$, so

$$F\colon M \to S^n.$$

Moreover, $F(0,x) = x$ and $F(1,x) = -x$. Let $\mathit{\Omega}$ be the solid angle form, and let $\omega = F^* \mathit{\Omega}$. Since d commutes with pullback of forms,

$$d\omega = F^* d\mathit{\Omega} = 0.$$

By Stokes's theorem,

$$\begin{aligned} 0 &= \int_M d\omega, \\ &= \int_{\partial M} \omega, \end{aligned}$$

and if we write $F_t(x)$ for $F(t,x)$, this is

$$= \int_{S^n} F_1^* \mathit{\Omega} - \int_{S^n} F_0^* \mathit{\Omega},$$

and since F_1 reverses orientation, we pop out a minus sign in the first term, but then it also affects each tangent plane by -1, so a determinant of $(-1)^n$ on those tangent planes, so

$$\begin{aligned} &= (-1)^{n+1} \int_{S^n} \mathit{\Omega} - \int_{S^n} \mathit{\Omega}, \\ &= \left((-1)^{n+1} - 1\right) \int_{S^n} \mathit{\Omega}. \end{aligned}$$

If n is even, this is

$$2 \int_{S^n} \mathit{\Omega} \neq 0,$$

a contradiction.

On the other hand, if n is odd, we can construct a nonwhere zero tangent vector field: $n+1$ is even and we can split $\mathbb{R}^{n+1}$ into a direct sum of mutually orthogonal planes:

$$\mathbb{R}^{n+1} = \mathbb{R}^2 \oplus \mathbb{R}^2 \oplus \cdots \oplus \mathbb{R}^2.$$

Take the linear map $T \colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ which just rotates each of these planes by a right angle, and let $X(x) = Tx$. □

Review problems

14.4 Denote by $\mathrm{GL}\,(n,\mathbb{R})$ the set of invertible $n \times n$ matrices. What is $T_I \mathrm{GL}\,(n,\mathbb{R})$?

14.5 Denote by $\mathrm{SL}\,(n,\mathbb{R})$ the set of $n \times n$ matrices of determinant 1. What is $T_I \mathrm{SL}\,(n,\mathbb{R})$?

14.6 Denote by $\mathrm{O}\,(n)$ the set of $n \times n$ orthogonal matrices. What is $T_I \mathrm{O}\,(n)$?

14.7 Denote by $\mathrm{SO}\,(n)$ the set of $n \times n$ rotation matrices (in other words: orthogonal matrices of determinant 1). What is $T_I \mathrm{SO}\,(n)$?

Notes

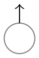
1. The problem of determining the maximum number of everywhere linearly independent vector fields on a sphere took many years to solve, and was important in the development of topology.
2. The coconut theorem is also called the *hairy ball theorem* or the *hedgehog theorem*

Chapter 15

Volumes of Manifolds

We explain how to define and calculate the volume of a manifold.

Normal vector fields

Why did we say that an orientable surface was one which has "two sides"? A *normal vector* to a manifold $M \subset \mathbb{R}^n$ at a point $x \in M$ is a vector $v \in \mathbb{R}^n$ which is perpendicular to every vector in T_xM.  A *normal vector field* on a manifold M is a continuous map $X\colon M \to \mathbb{R}^n$ so that $X(x)$ is a normal vector, for each $x \in M$. It is standard to write $S^{n-1} \subset \mathbb{R}^n$ to mean the unit sphere around the origin. (Careful: the superscript ${}^{n-1}$ means that the sphere has dimension $n-1$. It doesn't mean $(n-1)$-fold Cartesian product.) For example, on the unit sphere $S^n \subset \mathbb{R}^{n+1}$, the vector field $X(x) = x$ is a normal vector field. It is a *unit* normal vector field, i.e. $\|X\| = 1$. A manifold $M \subset \mathbb{R}^{n+1}$ of dimension n (in other words, of one less dimension than the ambient space) is called a *hypersurface*.

Theorem 15.1. *A hypersurface is orientable just when it admits a continuous nowhere vanishing normal vector field. This occurs just when it admits a continuously differentiable unit normal vector field, which is then unique up to* $\pm$.

Proof. Suppose that $M \subset \mathbb{R}^{n+1}$ is a hypersurface. Near each point of M, we can pick a chart, say h. Then the vector fields $X_i = h_* e_i$ are linearly independent. A collection of linearly independent vector fields forming a basis of T_xM at each point x is called a *framing*. So we know that near each point of M, there is a framing.

Once we have a framing, we can apply Gram–Schmidt orthogonalization to it to produce an orthonormal framing. So we can cover M by open sets on each of which we have an orthonormal framing, say $X_1, X_2, \dots, X_n$, by continuously differentiable vector fields.

Consider trying to pick a unit length normal vector ν at some point $x \in M$. It must be perpendicular to all of the vectors $X_1(x), X_2(x), \dots, X_n(x)$. So there are only 2 possible choices, since the orthogonal complement to these is a 1-dimensional vector space. There is a unique choice ν for which

$$\nu, X_1(x), X_2(x), \dots, X_n(x)$$

is a positively oriented orthonormal basis, i.e. a rotation of the standard basis.

Near each point of M, we can select some continuously differentiable vector field Y which is linearly independent of $X_1, X_2, \dots, X_n$, and then apply Gram–Schmidt orthogonalization to Y to obtain a unit normal vector field, say ν. If needed, replace ν by $-\nu$ to arrange that

$$\nu(x), X_1(x), X_2(x), \dots, X_n(x)$$

is a positively oriented orthonormal basis wherever defined. The orthogonalization process consists of arithmetic operations applied to the various vector fields, so the resulting $\nu(x)$ normal vector field is continuously differentiable.

To each orthonormal framing $X_1, X_2, \dots, X_n$ on some open subset of M, we have added a unit normal vector field ν. If we change the choice of orthonormal coframing by some rotation at each point, clearly by continuity we don't change the choice of ν. But if we change the choice by an orthogonal transformation with determinant -1, we then have to change the sign of ν to make the resulting basis $\nu, X_1, X_2, \dots, X_n$ be positively oriented. If we cover M in charts, the resulting normals will then agree on overlaps of charts just when the charts arecompatibly oriented.

If we start out with a choice of continuous unit normal vector field ν, then by the uniqueness above, we can select those charts which produce that same unit vector field via our procedure above. If a chart produces $-\nu$ via that procedure, then compose the chart with an orientation reversing diffeomorphism to get a new chart. $\square$

15.1 Suppose that $M \subset \mathbb{R}^{n+1}$ is the set of zeroes of a function:

$$M = \left\{ x \in \mathbb{R}^{n+1} \mid f(x) = 0 \right\}$$

where $f \colon \text{open } \subset \mathbb{R}^{n+1} \to \mathbb{R}$ is continuously differentiable and $df \neq 0$ at every point of M. Prove that M is a manifold. If we write

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i$$

then prove that the vector field

$$\nu = \frac{1}{\sqrt{\sum_i \left(\frac{\partial f}{\partial x_i}\right)^2}} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_{n+1}} \end{pmatrix}$$

is a unit normal vector field to M.

15.2 Use the result of problem 15.1 to find the unit normal vector field to the ellipsoid

$$\sum_i \left(\frac{x_i}{a_i}\right)^2 = 1$$

in $\mathbb{R}^n$, where $a_1, a_2, \dots, a_n > 0$ are constants.

15.3 Suppose that $M \subset \mathbb{R}^{n+1}$ is the graph of a function:

$$M = \left\{ \, (x, y) \in \mathbb{R}^{n+1} \;\middle|\; y = f(x) \, \right\}$$

where $f \colon \text{open } \subset \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. Prove that M is a manifold. If we write

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i$$

then prove that the vector field

$$\nu = \frac{1}{\sqrt{1 + \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2}} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \\ -1 \end{pmatrix}$$

is a unit normal vector field to M.

The volume form

In $\mathbb{R}^n$, we write $dV_{\mathbb{R}^n}$ to denote the *volume form*

$$dx_1 \wedge dx_2 \wedge \dots \wedge dx_n.$$

Careful: it is not d of any function V. We may also denote it dV

If M is an oriented hypersurface, say with unit normal vector field ν, then we define dV_M, the *volume form*, also often called the *area form* (especially if M is 2-dimensional) by

$$dV_M = \nu \lrcorner dV_{\mathbb{R}^{n+1}}.$$

If we call it the *area form*, then we denote it dA_M.

Define the *volume*, also called the *area*, of a hypersurface M to be

$$\int_M dV_M.$$

For example, on the unit sphere S^{n-1}, we saw the volume form already: $dV_{S^{n-1}} = \Omega$, the solid angle form restricted to the unit sphere. Our polar coordinate formula is thus

$$\int_{\mathbb{R}^n} f(x) \, dV = \int_0^\infty \left(\int_{S^{n-1}} f(ru) dV_{S^{n-1}} \right) r^{n-1} \, dr.$$

For example, a very important integral:

$$\int_{\mathbb{R}^n} e^{-\|x\|^2}\, dV = \operatorname{Area}\left(S^{n-1}\right) \int_0^\infty e^{-r^2} r^{n-1}\, dr.$$

Another way to calculate that integral: write each vector $x \in \mathbb{R}^n$ as $x = (s, t)$, for some $s \in \mathbb{R}^2$ and $t \in \mathbb{R}^{n-2}$. Clearly

$$\|x\|^2 = \|s\|^2 + \|t\|^2,$$

so

$$\int_{\mathbb{R}^n} e^{-\|x\|^2}\, dV = \int_{\mathbb{R}^2} e^{-\|s\|^2} \int_{\mathbb{R}^{n-2}} e^{-\|t\|^2}$$

giving a relationship between areas of spheres, which we can use inductively to find the area of a sphere in any dimension from the length of the unit circle. We leave the reader to prove:

$$\operatorname{Area}\left(S^n\right) = \begin{cases} \frac{(2\pi)^{n/2}}{2\cdot 4\cdot\dots\cdot(n-2)}, & \text{if } n \text{ is even,} \\ \frac{2(2\pi)^{(n-1)/2}}{1\cdot 3\cdot\dots\cdot(n-2)}, & \text{if } n \text{ is odd.} \end{cases}$$

15.4 Use the result of problem 15.2 on page 98 to find the volume form of the ellipsoid

$$\sum_i \left(\frac{x_i}{a_i}\right)^2 = 1$$

in $\mathbb{R}^n$, where $a_1, a_2, \dots, a_n > 0$ are constants.

How can we define the volume form on a manifold of any dimension, for example a curve in $\mathbb{R}^3$?

Lemma 15.2. *Suppose that M is a p-dimensional manifold of $\mathbb{R}^n$. Then M is orientable just when there is a p-form on M which doesn't vanish anywhere. If M is oriented, then M has a p-form, called the* volume form *of M, which is the unique nowhere vanishing p-form dV_M on M so that if $X_1, X_2, \dots, X_p$ is a positively oriented orthonormal framing, then*

$$dV_M\left(X_1, X_2, \dots, X_p\right) = 1.$$

Moreover $\int_M f\, dV_M > 0$ for any nonnegative continuous function f, somewhere positive, and vanishing outside a compact set. We define

$$\operatorname{Vol}(M) = \int_M dV_M$$

(which could be infinite).

Proof. Start by taking a positively oriented orthonormal framing $X_1, X_2, \dots, X_p$ in some open set on M. We want to see why there is a unique p-form ω so that

$$\omega\left(X_1, X_2, \dots, X_p\right) = 1.$$

Look in a chart, so without loss of generality just assume that $M \subset \mathbb{R}^p$ is an open subset (or $M \subset [0,1]^p$, if you have corners). Then look at any p-form:

$$\omega = f\, dx_1 \wedge dx_2 \wedge \dots \wedge dx_n.$$

Clearly if we want to solve the equation

$$\omega\left(X_1, X_2, \dots, X_p\right) = 1$$

then that completely determines the function f. So there is only one choice of p-form ω, which must be dV_M.

Pick a compatibly oriented chart h. Then applying Gram–Schmidt orthogonalization to $h_* e_i$, we get some positively oriented orthonormal framing X_i. We know that

$$dV_M\left(X_1, X_2, \dots, X_p\right) = 1.$$

But these X_i are some positive determinant linear combinations of the $h_* e_i$. Therefore

$$h^* dV_M = g\, dx_1 \wedge dx_2 \wedge dx_p$$

for some positive function $g(x)$. So if $f \ge 0$ is continuous and not everywhere 0, and $f = 0$ outside the image of the chart h, then

$$\int_M f\,\omega = \int f \circ h g\, dx_1 \wedge dx_2 \wedge dx_p > 0.$$

It then follows that all integrals of continuous nonnegative functions are nonnegative, and positive unless the function vanishes.

Suppose that I pick a positively oriented orthonormal framing $X_1, X_2, \dots, X_p$ on an open set U, and construct a p-form ω as above. Suppose that you pick a positively oriented orthonormal framing $\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_p$ on an open set $\tilde{U}$, and construct a p-form $\tilde{\omega}$ as above. Since your vector fields form a basis for each tangent space, as do mine,

$$\tilde{X}_i = \sum_j g_{ji} X_j,$$

for some functions g_{ji}. In fact, the matrix of functions $g = (g_{ij})$ is a rotation matrix, since both bases are positively oriented and orthonormal. Calculate:

$$\begin{aligned} 1 &= \tilde{\omega}\left(\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_p\right), \\ &= \sum_{j_1 j_2 \dots j_p} g_{1j_1} g_{2j_2} \dots g_{pj_p} \tilde{\omega}\left(X_{j_1}, X_{j_2}, \dots, X_{j_p}\right), \end{aligned}$$

and we let the reader carry out the linear algebra needed to prove that this is

$$= \det(g)\tilde{\omega}\left(X_1, X_2, \dots, X_p\right),$$

but $\det(g) = 1$, since g is a rotation, so

$$= \tilde{\omega}\left(X_1, X_2, \dots, X_p\right).$$

Therefore $\tilde{\omega} = \omega$. ☐

15.5 Prove that $\operatorname{Vol}(M \times N) = \operatorname{Vol}(M)\operatorname{Vol}(N)$.

To calculate volumes:

Lemma 15.3. *Suppose that $h\colon U \to M$ is a positively oriented chart for an oriented manifold M, perhaps with boundary and corners. If we write h as $x = h(u)$, say for $u \in U$, then*

$$h^* dV_M = \sqrt{\det G}\, dV_{\mathbb{R}^n}$$

where G is the matrix

$$G = \begin{pmatrix} \left\langle \frac{\partial x}{\partial u_1}, \frac{\partial x}{\partial u_1} \right\rangle & \left\langle \frac{\partial x}{\partial u_1}, \frac{\partial x}{\partial u_2} \right\rangle & \dots & \left\langle \frac{\partial x}{\partial u_1}, \frac{\partial x}{\partial u_n} \right\rangle \\ \left\langle \frac{\partial x}{\partial u_2}, \frac{\partial x}{\partial u_1} \right\rangle & \left\langle \frac{\partial x}{\partial u_2}, \frac{\partial x}{\partial u_2} \right\rangle & \dots & \left\langle \frac{\partial x}{\partial u_2}, \frac{\partial x}{\partial u_n} \right\rangle \\ \vdots & \vdots & \vdots & \vdots \\ \left\langle \frac{\partial x}{\partial u_n}, \frac{\partial x}{\partial u_1} \right\rangle & \left\langle \frac{\partial x}{\partial u_n}, \frac{\partial x}{\partial u_2} \right\rangle & \dots & \left\langle \frac{\partial x}{\partial u_n}, \frac{\partial x}{\partial u_n} \right\rangle \end{pmatrix}.$$

Proof. Let $Y_i = \frac{\partial h}{\partial u_i}$. So $Y_1, Y_2, \dots, Y_n$ is a basis for $T_m M$ at each $m \in M$, if M is n-dimensional. Apply the Gram–Schmidt process to get an orthonormal basis, say $X_1, X_2, \dots, X_n$, for $T_m M$. Somehow write the Y_i as a linear combination of the X_i:

$$Y_i = \sum_j a_{ij} X_j.$$

Let $A = (a_{ij})$, so

$$\omega\left(Y_1, Y_2, \dots, Y_n\right) = \det A.$$

Take inner products:

$$\begin{aligned} \left\langle Y_i, Y_j \right\rangle &= \left\langle \sum_k a_{ik} X_k, \sum_\ell a_{j\ell} X_\ell \right\rangle, \\ &= \sum_k a_{ik} a_{k\ell}, \\ &= \left(AA^t\right)_{ij}. \end{aligned}$$

So $G = AA^t$. In particular, we see that G is a postive definite symmetric matrix and $\det G = \det A \det A^t$. But $\det A^t = \det A$, so

$$\begin{aligned} \sqrt{\det G} &= \det A, \\ &= \omega\left(Y_1, Y_2, \dots, Y_n\right), \\ &= \omega\left(h_* e_1, h_* e_2, \dots, h_* e_n\right), \\ &= h^* \omega\left(e_1, e_2, \dots, e_n\right). \end{aligned}$$

□

15.6 Use this to find an integral expression for the length of the curve $y = f(x)$ in the plane $\mathbb{R}^2$ for $f \colon [a, b] \to \mathbb{R}$ a function, $a < b \in \mathbb{R}$.

For example, take as M any curve in the (x, y)-plane. Let

$$h(t) = (x(t), y(t))$$

be a chart. Then

$$\begin{aligned} G &= \left(\left\langle \tfrac{\partial(x,y)}{\partial t}, \tfrac{\partial(x,y)}{\partial t} \right\rangle\right), \\ &= \left(\frac{\partial x}{\partial t}\right)^2 + \left(\frac{\partial y}{\partial t}\right)^2, \end{aligned}$$

so that

$$h^* ds = \sqrt{\left(\frac{\partial x}{\partial t}\right)^2 + \left(\frac{\partial y}{\partial t}\right)^2}\, dt,$$

and the length of the curve is

$$\int \sqrt{\left(\frac{\partial x}{\partial t}\right)^2 + \left(\frac{\partial y}{\partial t}\right)^2}\, dt,$$

the usual formula for length of a plane curve.

Review problems

15.7 Suppose that $M \subset \mathbb{R}^{n+1}$ is an n-dimensional manifold. Let $M_1 \subset \mathbb{R}^{2n}$ be the set of pairs (x, v) where $x \in M$ and v is a unit length vector normal to M at x. Prove that M_1 is an n-dimensional manifold. Define an orientation on M_1. Prove that the map $(x, v) \in M_1 \mapsto x \in M$ is a 2-1 local diffeomorphism. If M is connected, prove that either (1) M is unorientable and M_1 is connected or (2) M is orientable and M_1 is not connected.

Notes

On a 1-dimensional manifold, we always write the volume form as ds and refer to it as the *arc length form.* Physicists often say *volume element*, *area element*, and *length element* instead of volume form, area form and arc length form.

Chapter 16

Sard's Theorem

You know that for max and min problems, we only need to look at critical points. We will see that for many problems in analysis, we only need to look away from critical points.

Regular and critical points

A *critical point* of a continuously differentiable map $f\colon P \to Q$ between manifolds is a point $p \in P$ where $f'(p)$ is not a surjective linear map; any other point is called a *regular point*. A point $q \in Q$ is called a *critical value* of a continuously differentiable map $f\colon P \to Q$ if $q = f(p)$ for some critical point $p \in P$. A point $q \in Q$ which is not a critical value is called a *regular value*. In other words, a point $q \in Q$ is a regular value if there is no critical point $p \in P$ so that $f(p) = q$.

16.1 Which are the regular points, critical points, regular values and critical values of

1. $y = x^2$,
2. $y = x^3 - x^2$,
3. $y = 1$,
4. $y = \sin x$,
5. $y = \arctan x$?

Theorem 16.1 (The implicit function theorem)**.** *Suppose that $P \subset \mathbb{R}^n$ is a p-dimensional manifold, $Q \subset \mathbb{R}^m$ is a q-dimensional manifold, and $f\colon P \to Q$ is a continuously differentiable map. For each regular value $q_0 \in Q$, the set $M = f^{-1}\left\{q_0\right\}$ is a $(p-q)$-dimensional manifold in $\mathbb{R}^n$ with tangent space*

$$T_p M = \operatorname{kernel} f'(p)$$

for every point $p \in M$.

16.2 Generalize this implicit function theorem to an analogue of the rank theorem.

Measure zero

A set of points $A \subset \mathbb{R}^n$ has *measure zero* if for any $\varepsilon > 0$ there is a sequence of balls covering A whose volumes sum to less than ε.

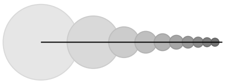
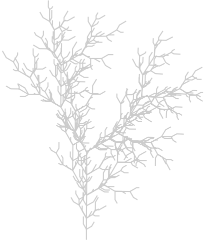


Figure 16.1: The real number line is a measure zero subset of the plane.

16.3 Prove that a line is a measure zero subset of the plane.

16.4 Prove that we could use boxes instead of balls in the definition of measure zero.

16.5 Use the previous problem to prove that we could use arbitrarily rotated boxes instead of balls in the definition of measure zero.



A measure zero set might look like this

Lemma 16.2. *The union of any sequence of measure zero sets has measure zero.*

Proof. Choose some $\varepsilon > 0$ and then pick $\varepsilon_1, \varepsilon_2, \dots \to 0$ positive so that

$$\sum_i \varepsilon_i \le \varepsilon.$$

□

Lemma 16.3. *The image of a measure zero set under a continuously differentiable map $f\colon$ open $\subset \mathbb{R}^n \to \mathbb{R}^n$ has measure zero.*

Proof. We can cover the domain of f in a sequence of closed balls, say all of the balls with rational radii and center points, and it is enough to prove the result for each of these, so we can assume that $f\colon \bar{B} \to \mathbb{R}^n$ is defined on a closed ball $\bar{B} \subset \mathbb{R}^n$. By continuity we can bound $\|f'(x)\|$ on $\bar{B}$, say $\|f'(x)\| < C$. By theorem 3.7 on page 26, we can then say that f moves points apart by a factor at worst C, i.e. f takes each ball, say of radius r, inside a ball of radius at most Cr. If we take a measure zero set $A \subset \bar{B}$ and make a covering of A by balls, say of radii r_i, so that their sum of volume is small, say less than ε/C^n, then we can cover $f(A)$ by balls whose volume sum is not more than ε. □

Lemma 16.4. *Suppose that $U \subset \mathbb{R}^n$ is an open set. A subset $A \subset U$ has measure zero just when, for any open cover of U, say by sets W_α, $A \cap W_\alpha$ has measure 0 for all α.*

Proof. Clearly if A has measure zero, then $A \cap W$ does, for any open set W. Conversely, suppose that $A \cap W_\alpha = 0$ for all α. Let $A_j = A \cap \bar{B}_{r_j}(0)$ for some sequence of radii $r_j \to \infty$. We need only prove that A_j has measure zero, for all j. Fix some value of j. By compactness, we can find a finite cover of $\bar{B}_{r_j}(0)$, say by $W_{\alpha_1}, W_{\alpha_2}, \dots, W_{\alpha_k}$. Then

$$\begin{aligned} A_j &= (A_j \cap W_{\alpha_1}) \cup (A_j \cap W_{\alpha_2}) \cup \dots \cup (A_j \cap W_{\alpha_k}), \\ &\subset (A \cap W_{\alpha_1}) \cup (A \cap W_{\alpha_2}) \cup \dots \cup (A \cap W_{\alpha_k}) \end{aligned}$$

has measure zero, since it is contained in a finite union of measure zero sets. □

Lemma 16.5. *If a box X can be covered by a collection of boxes $X_1, X_2, \dots, X_N$, then*

$$\operatorname{Vol}(X) \le \sum_i \operatorname{Vol}(X_i).$$

In particular, a box has measure zero just when one of its sides has length zero.

Proof. This is surprisingly difficult to prove. Suppose that $\operatorname{Vol}(X) > 0$. The volume of a box X is roughly equal to the number $N(X)$ of integer points lying in it. For example, the length of an interval $[a, b]$ is $b - a$, while the number $N([a, b])$ of integer points in the interval $[a, b]$ is between the limits

$$\lfloor b - a \rfloor \le N([a, b]) \le \lceil b - a \rceil.$$

Similarly, if

$$X = [a_1, b_1] \times [a_2, b_2] \times \dots [a_n, b_n]$$

then

$$\prod_j \lfloor b_j - a_j \rfloor \le N(X) \le \prod_j \lceil b_j - a_j \rceil.$$

Suppose that

$$X_i = [a_{i1}, b_{i1}] \times [a_{i2}, b_{i2}] \times \dots [a_{in}, b_{in}].$$

The integer points in X all lie in at least one of $X_1, X_2, \dots, X_N$, so

$$N(X) \le \sum_i N(X_i),$$

and so

$$\prod_j \lfloor b_j - a_j \rfloor \le \sum_i \prod_j \lceil b_{ij} - a_{ij} \rceil,$$

If $\lambda > 0$ is any number, let λX be the rescaled box, so that $x \in X$ just when $\lambda x \in \lambda X$. Then

$$N(\lambda X) \le \sum_i N(\lambda X_i),$$

so as before

$$\prod_j \lfloor \lambda b_j - \lambda a_j \rfloor \le \sum_i \prod_j \lceil \lambda b_{ij} - \lambda a_{ij} \rceil.$$

We can estimate these for large λ by

$$\prod_j (\lambda b_j - \lambda a_j - 1) \le \sum_i \prod_j (\lambda b_{ij} - \lambda a_{ij} + 1).$$

Divide both sides by λ^n to get

$$\prod_j \left(b_j - a_j - \frac{1}{\lambda} \right) \le \sum_i \prod_j \left(b_{ij} - a_{ij} + \frac{1}{\lambda} \right).$$

Take the limit as $\lambda \to \infty$:

$$\text{Vol}\,(X) \le \sum_i \text{Vol}\,(X_i)\,.$$

If we can cover a box X by balls whose volumes sum to a small value ε, then we can cover by slightly larger boxes containing those balls, changing the volume sum to $2^n\varepsilon/c$, where c is the volume of the unit ball. Since n is fixed, we can make this as small as we like. So X has measure zero just when $\text{Vol}\,(X) = 0$. ☐

A set of points $A \subset M$ on a manifold M has *measure zero* for that manifold if for any chart h for M, the set $h^{-1}A$ has measure zero. By the above lemmas, this definition is independent of the choice of charts.

Lemma 16.6. *The empty set is the only measure zero open subset of any manifold.*

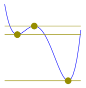
Proof. Since every ball in $\mathbb{R}^n$ contains a box of positive volume, no ball has measure zero. Every nonempty open subset of $\mathbb{R}^n$ is a union of a sequence of open balls, so does not have measure zero. The image of a chart is an open set, so does not have measure zero. ☐

Sard's theorem

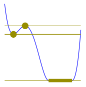
Consider one function of one variable $y = f(x)$. At a critical point, the rate of change is 0, so small changes in x make almost no change in y: a small interval of x values is mapped to a very small interval of y values.

Theorem 16.7 (Sard's theorem)**.** *Suppose that P and Q are manifolds of the same dimension and $f\colon P \to Q$ is a continuously differentiable map. The set of critical values of f has measure zero; in particular, f has regular values.*

For example, here there are only 3 critical values (3 heights where critical points occur):



In a similar example, you could have a long interval of critical points:



Here you still only have 3 critical values—the set of critical points might *not* have measure zero, but the set of critical values does.

Proof. By the above results, it is enough to work in a single chart on P and a single chart on Q, so we can assume that P and Q are bounded open subsets of $\mathbb{R}^n$. For simplicity, we can assume that P lies inside a box in $\mathbb{R}^n$, say of volume V, or even that P is such a box. Let

$$M = \max_{x \in P} \|f'(x)\|;$$

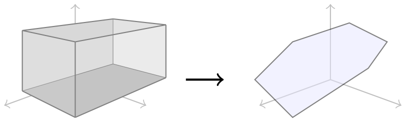
so f stretches distances by at most M.

Cover P in a grid. For any ε, if we make the grid fine enough, say of mesh size at most δ, then we can ensure that for any two points x and y in the same grid box,

$$\|f(y) - f(x) - f'(x)(y - x)\| < \varepsilon \|x - y\| < \varepsilon\delta. \tag{16.1}$$

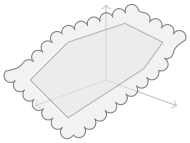
If there is no critical point of f inside a particular grid box, we just ignore that box. If there is a critical point of f inside a particular grid box, we take x to be such a point; so $f'(x)$ is not an invertible matrix.

Pick one such grid box, say X; we want to estimate $\operatorname{Vol}(f(X))$. As usual we can arrange $x = 0$ and $f(0) = 0$ by translation without changing the volume of $f(X)$. The image of $f'(0)$ lies in a linear subspace $W \subset \mathbb{R}^n$ of dimension at most $n - 1$, since the image of $f'(0)$ is not all of $\mathbb{R}^n$.



grid box X $\qquad$ $f'(0)X$

By equation 16.1, the set $f(X)$ lies within distance at most $\varepsilon\delta$ of $f'(0)X$, so $f(X)$ might look like a very slight thickening of $f'(0)X$:



$f(X)$

Rotation doesn't affect volume; rotate the picture to arrange that $W = \{0\} \times \mathbb{R}^{n-1}$. For the duration of this proof, let's adopt a notation in which X_r means a box all of whose sides have length r, centered at an unspecified point, and $\bar{B}_r$ similarly means a ball of radius r around an unspecified point. So if we write $B_r \subset X_{2r}$, we mean that every ball of radius r lies inside a box all of whose sides have length at most $2r$. Our box X is $X = X_\delta \subset B_{\delta\sqrt{n}/2}$, so

$$f'(0)X_\delta \subset f'(0)B_{\delta\sqrt{n}/2},$$

and $f'(x)$ stretches by at most a factor of M:

$$f'(0)X_\delta \subset \{0\} \times B_{M\delta\sqrt{n}/2} \subset \{0\} \times X_{M\delta\sqrt{n}}.$$

But $f\left(X_\delta\right)$ lies within $\varepsilon\delta$ of $f'(0)X_\delta$, so

$$f\left(X_\delta\right) \subset [-\varepsilon\delta, \varepsilon\delta] \times X_{M\delta\sqrt{n}+2\varepsilon\delta},$$

and therefore $f\left(X_\delta\right)$ lies in box of volume at most

$$2\varepsilon\delta \operatorname{Vol}\left(X_{M\delta\sqrt{n}+2\varepsilon\delta}\right) = 2\varepsilon\delta^n \left(M\sqrt{n}+2\varepsilon\right)^{n-1}.$$

The factor of ε is crucial—it describes the thinness of the image $f(X)$ of the grid box X and it doesn't show up if we work near a regular point.

Adding up over all of the grid boxes, each of which we can assume has volume δ^n, the critical values lie in various rotated boxes whose total volume is at most

$$2\varepsilon V \left(M\sqrt{n}+2\varepsilon\right)^{n-1},$$

which we can make arbitrarily small by choosing δ small enough to make ε as small as we need. □

Review problems

16.6 Prove that the set of rational numbers has measure zero in $\mathbb{R}$.

16.7 Suppose that $f_1, f_2, f_3, \dots$ is a sequence of functions $\mathbb{R} \to \mathbb{R}$. Prove that there is a value $y_0 \in \mathbb{R}$ so that none of these functions has y_0 as a critical value.

16.8 Suppose that $f \colon P \to Q$ is a continuously differentiable map of manifolds and that $\dim P < \dim Q$. Prove that the image of f has measure zero in Q.

Notes

1. The word “measure” just means “volume”, but it is tradition to say “measure zero” instead of saying “no volume”.
2. We have only given a special case of Sard's theorem following Spivak [7]; the general case is stated and proven in Sternberg [8].

Chapter 17

Homotopy and Degree

We think about continuous deformation of maps and we count how many times a point y is struck by a map $y = f(x)$. We explain how integrals change under some maps that are more complicated than diffeomorphisms.

Homotopy

Pick manifolds $P \subset \mathbb{R}^m$ and $Q \subset \mathbb{R}^n$. A *homotopy* between two continuous maps $f \colon P \to Q$ and $g \colon P \to Q$ is a continuous map

$$F \colon [0,1] \times P \to Q$$

so that $F(0,p) = f(p)$ and $F(1,p) = g(p)$ for all $p \in P$. Write $F_t(p)$ to mean $F(t,p)$. So F_t is a "deformation" of f into g.

Degree

Suppose that $f \colon P \to Q$ is a continuously differentiable map between two oriented manifolds of equal dimension. The *sign* of f at a regular point $p \in P$ is 1 if f is orientation preserving near p, and -1 is f is orientation reversing near p. Suppose that $q_0 \in Q$ is a regular value and that there are only finitely points $p \in P$ mapped to q_0 by f. The *degree* of f above q_0 is the sum of signs at these points p:

$$\deg_{q_0} f = \sum_{f(p)=q_0} (\text{sign of } f \text{ at } p)\,.$$

Proper maps

A map $f \colon P \to Q$ between manifolds is *proper* if every compact set $K \subset Q$ has compact preimage $f^{-1}K \subset P$. If M is a p-dimensional manifold with boundary and corners, the *corners of codimension* k are the connected subsets of ∂M which are identified, in some chart, with intersections of k walls. A map $f \colon P \to Q$ between manifolds with boundary and corners is *neat* if f takes corners of any codimension in P to corners of the same codimension in Q; if P has no corners or boundary then by default any map on P is neat.

17.1 Prove that every continuous map from a compact manifold without boundary is proper.

17.2 Which polynomial functions $y = f(x)$ of one variable are proper?

Suppose that $f: P \to Q$ is a continuously differentiable neat map on a manifold P with boundary and corners. A point $p \in P$ is *regular* for f if f' is surjective as a linear map $T_pP \to T_qQ$ and also surjective as a linear map $T_pC \to T_qD$ for any corner C of P containing p and being mapped to some corner D of Q of equal codimension. A regular value is a point of Q whose preimage consists only in regular points.

Theorem 17.1 (The implicit function theorem)**.** *Suppose that $P \subset \mathbb{R}^n$ is a p-dimensional manifold, perhaps with boundary and corners, $Q \subset \mathbb{R}^m$ is a q-dimensional manifold, perhaps with boundary and corners. Suppose that $f: P \to Q$ is a neat continuously differentiable map on a manifold P with boundary and corners. For each regular value $q_0 \in Q$, the set $M = f^{-1}\left\{q_0\right\}$ is a $(p-q)$-dimensional manifold in $\mathbb{R}^n$, perhaps with boundary and corners, with tangent space*

$$T_pM = \operatorname{kernel} f'(p)$$

for every point $p \in M$, and and the corners of M of codimension k all have the form $C \cap M$ where C is a codimension k corner of P.

The proof of the implicit function theorem “with corners” is the same as the earlier proof without corners.

Lemma 17.2. *For any regular value q_0 of any continuously differentiable neat and proper map $f: P \to Q$ between manifolds of equal dimension, there are only finitely many points $p \in P$ mapping to q_0.*

Proof. If we have an infinite sequence of such points, all distinct, then they all lie inside the compact set $f^{-1}(q_0)$. Some subsequence must have a limit, say p. But then $f(p) = q_0$ is a regular value, so f is a diffeomorphism near p, and therefore no point close to p maps to q_0 other than p. $\square$

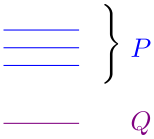
Lemma 17.3. *Roughly: small sets have small preimages under proper maps. More precisely: if $f: P \to Q$ is a proper map, and $K_Q \subset Q$ is a compact set and $K_P = f^{-1}K_Q$, then for every open neighborhood U_P of K_P, there is an open neighborhood U_Q of K_Q so that $f^{-1}U_Q \subset U_P$.*

Proof. Take a sequence $W_1, W_2, \dots$ of every smaller neighborhoods of K_Q, so that $\bar{W}_{i+1} \subset W_i$ and so that $\cap_i W_i = K_Q$. For example, you could take W_i to be all points of Q of distance less than ε_i from K_Q, as long as the sequence ε_i decreases and $\varepsilon_i \to 0$. We can further assume that all of the sets W_i are contained in a single compact set $\bar{W}_1$. We want to prove that we can use one of these W_i as our choice of U_Q. If not, then each $f^{-1}W_i$ must not lie in U_P, so contains a point, say $x_i \notin U_P$. These points x_i all lie in the compact set $f^{-1}\bar{W}_i$.

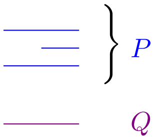
So there must be a convergent subsequence. We can drop all of the elements of the sequence except those in the subsequence, and relabel, to arrange that the sequence converges, say $x_i \to x$. So $f(x) \in \bar{W}_i$ for all i, and therefore $f(x) \in K_Q$, so $x \in f^{-1}K_P \subset U_P$, so $x \in U_P$. The set U_P is open, so all points sufficiently close to x lie in U_P. In particular, all but finitely many of the x_i lie in U_P, a contradiction. $\square$

17.3 How small does U_Q have to be? Take the map $y = x^k$ and $K_Q = (0)$. Then $K_P = (0)$, and we can try $U_P = (-\varepsilon, \varepsilon)$. What choices of U_Q can you make?

Imagine a map $f \colon P \to Q$ as a projection of a surface P to its shadow in a plane Q. "Above" each of those regular points in Q, there are various layers of P, known as *sheets.*



We have a local inverse map for each sheet. It is clear then that the degree of a proper map cannot change as we move along a curve in Q, as long as that curve only passes through regular points. If we have an improper map, the sheets up in P might "run out":



Then we don't have an inverse map defined for that sheet.

Lemma 17.4. *Suppose that $f \colon P \to Q$ is a neat and proper continuously differentiable map between manifolds of the same dimension. Then the set of regular values of f is open. Moreover, each regular value lies in an open set W so that $f^{-1}W$ can be written as a union*

$$f^{-1}W = U_1 \cup U_2 \cup \cdots \cup U_m,$$

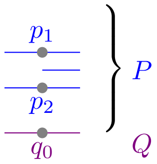
where these U_i are disjoint open sets and

$$f|_{U_i} : U_i \to W$$

is a diffeomorphism.

In a picture, P "above" the open set $W \subset Q$ looks like a collection of "sheets" $U_1, U_2, \dots, U_m$, all of which are diffeomorphic copies of W: the sheets don't run out.

Proof. Suppose that $q_0 \in Q$ is a regular value. Suppose that the points $p \in P$ at which $f(p) = q_0$ are $p = p_1, p_2, \dots, p_m$.



By the inverse function theorem, around each of these points p_i, there is some open set $U_i \subset P$ and there is some open set $W_i \subset Q$ so that

$$f|_{U_i} : U_i \to W_i$$

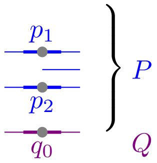
is a diffeomorphism. Let

$$W = W_1 \cap W_2 \cap \dots \cap W_m.$$

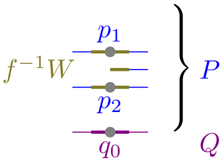
Replace U_i by $U_i \cap f^{-1}W$ so that

$$f|_{U_i} : U_i \to W$$

is a diffeomorphism for each i.



In our picture, W is the thick purple line, U_1 and U_2 are the thick blue lines. The preimage $f^{-1}W$ consists of U_1, U_2, and a little bit of the middle blue line:



Let $K_Q = \{q_0\}$. The set $U_P = U_1 \cup U_2 \cup \dots \cup U_m$ is an open neighborhood of the set

$$K_P = f^{-1}K_Q = \{p_1, p_2, \dots, p_m\}.$$

By lemma 17.3 on page 112, there is an open set U_Q containing q_0 so that $f^{-1}U_Q$ lies inside U_P. Replace W by $W \cap U_Q$, and replace each U_i by $f^{-1}W \cap U_i$, and you get $f^{-1}W = U_1 \cup U_2 \cup \dots \cup U_m$. So our picture was wrong: the middle blue line we drew can't really be there. $\square$

17.4 Suppose that $f\colon P \to Q$ is a proper continuously differentiable map, and P and Q are manifolds of equal dimension, and this dimension is 2 or more, and that Q is connected, and f has only finitely many critical points in any compact subset of P. Prove that the set of regular values is connected, and that any two regular values have the same (possibly infinite) number of preimages. Prove that f is surjective. Use this to prove that any nonconstant complex polynomial function of a complex variable has a root.

Lemma 17.5. *Suppose that $f\colon P \to Q$ and $g\colon P \to Q$ are neat and proper continuously differentiable maps between oriented manifolds of equal dimension, perhaps with boundary and corners. Suppose that P is compact and that there is a neat and proper continuously differentiable homotopy between f and g. Suppose that $q_0 \in Q$ is a regular point for f and for g. Then*

$$\deg_{q_0} f = \deg_{q_0} g.$$

Proof. Take a neat and proper continuously differentiable homotopy F between f and g. Suppose first that q_0 is a regular value for F. By the implicit function theorem, $F^{-1}(q_0) \subset [0,1] \times P$ is a 1-dimensional manifold, with boundary:

$$\partial F^{-1}\{q_0\} = f^{-1}\{q_0\} \times \{0\} \cup g^{-1}\{q_0\} \times \{1\}.$$

This manifold is oriented, using the orientations for $[0,1]$, P and Q. Integrate 1, by Stokes's theorem:

$$\begin{aligned} 0 &= \int_{F^{-1}(q_0)} d1, \\ &= \int_{g^{-1}(q_0)} 1 - \int_{f^{-1}(q_0)} 1, \\ &= \deg_{q_0} g - \deg_{q_0} f. \end{aligned}$$

Suppose that q_0 is not a regular value of F. Since q_0 is regular for f, by lemma 17.4 on page 113, every point close enough to q_0 is regular for f. Similarly all points close enough to q_0 are regular for g. The set of critical points of F cannot contain an open set by Sard's theorem and lemma 16.6 on page 108. Therefore F and f and g are all regular at points arbitrarily close to q_0, where our argument above tells us that f and g have the same degree. The degrees of f and g are therefore equal at all of those points. But the degree of f doesn't change near q_0, since f is regular at q_0, and the same is true for g. $\square$

Flows of vector fields

Besides the existence and uniqueness of solutions of ordinary differential equations, you also know about the continuous dependence on initial conditions:

Theorem 17.6 (Picard [1] p. 224)**.** *Every continuously differentiable vector field X on an open set $U \subset \mathbb{R}^n$. There is a unique continuously differentiable* *map $F\colon \text{open} \subset \mathbb{R} \times U \to U$, called the* flow *of the vector field X, so that for any chosen point $p \in U$, the map $t \mapsto F(t,p)$ is the flow line of X through p. For any chosen time t, the map $F_t\colon \text{open} \subset U \to \text{open} \subset U$, given by $F_t(x) = F(t,x)$, is a diffeomorphism where defined.*

By compactness, if X vanishes outside a compact set $K \subset U$, then F is defined on all of $\mathbb{R} \times U$. By uniqueness, $F(t,x) = x$ for any $x \in U - K$.

The diffeomorphism group

Lemma 17.7. *For any point x_0 inside a closed ball $\bar{B} \subset \mathbb{R}^n$, not on the boundary, there is an orientation preserving diffeomorphism $F\colon \bar{B} \to \bar{B}$ so that $F(0) = x_0$ and $F(x) = x$ for all points x near $\partial \bar{B}$.*

Proof. Take a bump function h in $\bar{B}$ vanishing near $\partial \bar{B}$ and equal to 1 in some smaller ball containing both x_0 and 0. Let X be the vector field $X(x) = h(x)x_0$. Since this vector field vanishes outside a compact set, its flow lines are defined for all time. Let $F(x)$ be the point that x gets to along the flow of X at time $t = 1$; F is a diffeomorphism, with inverse given by flowing along $-X$ for time $t = 1$. Inside the smaller ball $F(x) = x + x_0$. $\square$

A homotopy $F_t\colon P \to Q$ which is a diffeomorphism for all values of t is called an *isotopy.* Clearly F_0 is orientation preserving just when F_1 is.

A manifold M is *connected* if, for any two points $m_0 \in M$ and $m_1 \, inM$, there is a continuous path $m(t)$ from m_0 to m_1.

Lemma 17.8. *For any two points $m_0, m_1 \in M$ of a connected manifold (perhaps with boundary or corners), neither point contained in the boundary, there is an isotopy $F\colon [0,1] \times M \to M$ between the identity map and a map F_1, so that $F_1(m_0) = m_1$, and so that, for all $m \in M$ outside some compact set, $F_t(m) = m$ for $0 \le t \le 1$.*

Proof. Take a path from m_0 to m_1. Cover the path by images of charts, each chart mapping from a ball and taking the center of the ball to a point of the path. Call the image of this chart a “ball”, and the image of its center the “center” of that “ball”. The path is a map from a compact interval, so has compact image. Take a finite subcover: a finite set of charts. By lemma 17.7, we can make an isotopy of M which is the identity outside some compact set and which moves the “center” of the path along to some other point in the next “ball”. We can then make an isotopy to get that point to the “center” of that ball. After finitely many steps, we “glue” these isotopies together. The result is continuous, and if we reparameterize appropriately we can make it continuously differentiable, using the method of problem 2.12 on page 15. $\square$

17.5 You pick m distinct points of a manifold M, say $p_1, p_2, \dots, p_m$, and I pick another m distinct points $q_1, q_2, \dots, q_m$ of the same manifold. Prove that there is a diffeomorphism $f\colon M \to M$ so that for all $j = 1, 2, \dots, m$, $f(p_j) = q_j$.

Degree invariance

Theorem 17.9. *Consider a neat and proper continuously differentiable map from an oriented manifold (perhaps with boundary or corners) to a connected oriented manifold of the same dimension. The degree of this map is the same at all regular values.*

We can therefore replace the notation $\deg_{q_0} f$ by the notation $\deg f$.

Proof. Take a proper continuously differentiable map $f\colon P \to Q$ between two manifolds P and Q of equal dimension. Pick two regular points $q_0 \in Q$ and $q_1 \in Q$ not in the boundary. By lemma 17.8 on the preceding page, we can make an isotopy $F_t\colon Q \to Q$ so that $F_0(q) = q$ for all x, and $F_1(q_0) = q_1$, and so that outside some compact set K, $F_t(q) = q$ for all t and q. Look at the maps $F_t \circ f$. They are clearly proper. Moreover, it is clear that

$$\deg_{q_0} f = \deg_{F_t(q_0)} F_t \circ f$$

for any t, since these are counting the same points of P, and the same signs. By lemma 17.5 on page 115,

$$\deg_{q_1} F_1 \circ f = \deg_{q_1} f,$$

since these are homotopic maps. □

Integration

Theorem 17.10. *Suppose that $f\colon P \to Q$ is a neat and proper continuously differentiable map between two oriented manifolds (perhaps with boundary and corners) of the same dimension, say n. Suppose that Q is connected. For any n-form ω on Q, which vanishes outside some compact set,*

$$\int_P f^*\omega = \deg f \int_Q \omega.$$

Proof. If ω is supported in a small enough open set around a regular value, then the result is clear by lemma 17.4 on page 113. Using a partition of unity, the result becomes clear if ω vanishes near the critical values. Using the same partition of unity, we only need to prove the result for every n-form ω supported in arbitrarily small neighborhood of each critical value.

Pick a critical value $q_0 \in Q$ away from ∂Q. We leave the reader to prove that there is a continuously differentiable proper homotopy $F\colon [0,1] \times P \to Q$ so that $F_0 = f$ and so that F_1 has q_0 as a regular value, and $F_t = f$ on ∂P for all t. Pick an n-form ω on Q with compact support contained in the regular values of F_1. Because $d\omega = 0$,

$$\begin{aligned} 0 &= \int_{[0,1]\times P} F^* d\omega, \\ &= \int_{[0,1]\times P} dF^* \omega, \end{aligned}$$

to which we apply Stokes's theorem:

$$= \int_{\{1\}\times P} F^*\omega - \int_{\{0\}\times P} F^*\omega - \int_{[0,1]\times \partial P} F^*\omega,$$

but on $[0,1] \times \partial P$, $F^*\omega = f^*\omega$ is independent of t, so has no dt component, so the last term above vanishes

$$\begin{aligned} &= \int_P F_1^*\omega - \int_P f^*\omega, \\ &= \deg F_1 \int_Q \omega - \int_P f^*\omega, \\ &= \deg f \int_Q \omega - \int_P f^*\omega. \end{aligned}$$

By the same partition of unity argument, our result now holds for any n-form ω whose support stays away from ∂Q. Pick a continuous function h on Q, vanishing in a thin layer near ∂Q, and increasing to 1 on slightly thicker layer near ∂Q. By lemma 17.3 on page 112, if we make these layers thin enough, their preimages will be as thin as we want, lying close to ∂P. Therefore we can arrange that $\int_Q h\omega$ is as close to $\int_Q \omega$ as we want, and simultaneously that $\int_P f^* h\omega$ is as close to $\int_P f^*\omega$ as we want. By the above,

$$\int_P f^* h\omega = \deg f \int_Q h\omega.$$

Taking the limit as our layers get thin,

$$\int_P f^*\omega = \deg f \int_Q \omega.$$

□

In terms of functions instead of forms,

$$\int_X (f \circ F) \det F' = \deg F \int_{F(X)} f,$$

for any continuously differentiable proper map $F\colon X \to \mathbb{R}^n$, open set $X \subset \mathbb{R}^n$ and continuous function $f\colon F(X) \to \mathbb{R}$ vanishing outside a compact set.

Corollary 17.11. *If $f\colon P \to Q$ is a neat and proper continuously differentiable map between compact oriented manifolds of equal dimension, then we can write the degree as an integral:*

$$\deg f = \frac{\int_P f^* dV_Q}{\operatorname{Vol}(Q)}.$$

17.6 Consider the map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$,

$$(x, y) = f(s, t) = \left(\frac{s}{\sqrt{1+t^2}}, \frac{st}{\sqrt{1+t^2}} \right).$$

Prove that every disk around the origin in the x, y-plane has infinite area preimage in the s, t-plane. So there is no easy change of variables formula for integrals for improper maps.

The Gauss map

Suppose that $P \subset \mathbb{R}^3$ is an oriented surface. Let ν be the positively oriented unit normal vector. At each point $p \in P$, $\nu(p)$ is a unit length vector, so lies in the sphere $S^2 \subset \mathbb{R}^3$. The *Gauss map* is the map

$$G\colon P \to S^2, G(p) = \nu(p).$$

Careful: of course G is just ν, but we think of G as a map to a different manifold, while we think of ν as a normal vector field.

17.7 If P is a continuously differentiable manifold, prove that G is continuous.

Every 2-form on P can be written uniquely as $\omega = f\, dV_P$, for some function f. If P is a twice continuously differentiable manifold, then G is continuously differentiable, so we can write

$$G^* dV_{S^2} = K\, dV_P,$$

for a unique continuous function K, called the *Gauss curvature* of P. The Gauss curvature is large if the normal vector field changes rapidly in different directions, so that the Gauss map covers a large area of the sphere from a small area of P.

By corollary 17.11, we have the *Gauss–Bonnet theorem*:

$$\deg G = \frac{\int_P K\, dV}{4\pi} :$$

the total curvature added up over the surface (and suitably scaled) is an integer, representing the degree of the Gauss map. For example, if P is a sphere centered at the origin, then

$$G(x) = \nu(x) = \frac{x}{\|x\|}$$

has degree 1.

17.8 The same ideas work for curves in the plane. Define the Gauss map of a plane curve. Draw pictures to explain what the Gauss map of a figure 8 looks like. Find the degree of the Gauss map of a figure 8.

Notes

We have seen that the group of diffeomorphisms of any connected manifold M acts transitively; indeed we see that this is already the case for the group of all diffeomorphisms $f\colon M \to M$ which are equal to the identity outside some compact set. The diffeomorphism group is a mystery even for some well known manifolds.

Hints

1.11. The inverse is explicitly expressed as

$$A^{-1} = \frac{1}{\det A} \operatorname{adj}(A),$$

where $\operatorname{adj}(A)$ is the classical adjoint.
1.19. Let $\bar{A}$ be the intersection of all closed sets containing A. We need only show that $\bar{A}$ is closed; its complement is the union of all open sets not intersecting A, hence open.
2.1. Clearly $Ah = Bh + o(h)$. Take a vector $v \neq 0$ and let $h = tv$, for some small number t. $Atv = Btv + o(t)$. Divide by t, and send $t \to 0$: $Av = Bv$. Since this holds for any vector $v \neq 0$, it holds by linearity for all vectors v, so $A = B$.
2.2.

$$f'(x) = \begin{pmatrix} 1 & 1 \\ 1 & 2x_2 \end{pmatrix}.$$

2.4.

$$f'(x, y) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

2.5. If

$$f(x+h) - f(x) - f'(x)h = o(h),$$

then clearly as $h \to 0$,

$$|f(x+h) - f(x)| = |f'(x)h + o(h)| \to 0.$$

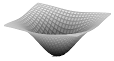
2.6. $f'(M)h = -M^{-1}hM^{-1}$
2.7. Write

$$f(x+h, y+k) - f(x, y) = (f(x+h, y+k) - f(x, y+k)) + (f(x, y+k) - f(x, y)).$$

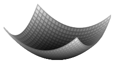
Pick k small enough to make the 2nd term as small as you need, and then pick h small enough to make the 1st term as small as you need.
2.10. Take the function f from problem 2.9 on page 15, or any other function which is continuously differentiable, positive for $|x| < 1$, and vanishing for $|x| \geq 1$. Let

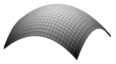
$$g(x) = \frac{\int_0^x f}{\int_0^1 f}.$$

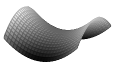
2.20.

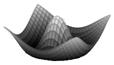

$-e^{-x^2-y^2}$

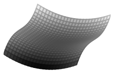
1

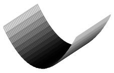

x^2+y^2

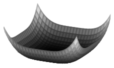

$-x^2-y^2$

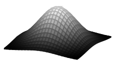

x^2-y^2

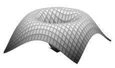

$-\left(x^2+y^2\right)e^{-x^2-y^2}$


x^2+y^3


x^2


x^4+y^4


$e^{-x^2-y^2}$

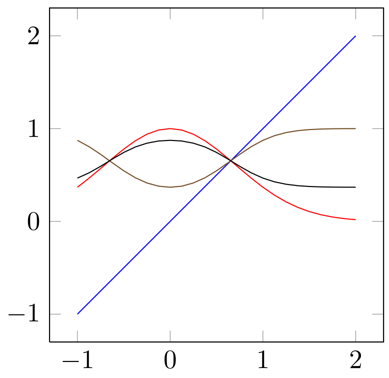

$\left(x^2+y^2\right)e^{-x^2-y^2}$

2.21. If $\alpha = 0$ or $\alpha > 1$, then $f'(0,0) = 0$. Otherwise, $f(x,y)$ is not differentiable at $(x,y) = (0,0)$.

2.24.

$$f(x,y) = \begin{cases} 0 & \text{if } x \le 0, \\ e^{-1/x^2} & \text{if } x > 0 \text{ and } y > 0, \\ e^{-1/x^4} & \text{if } x > 0 \text{ and } y < 0, \\ \text{undefined} & \text{if } x > 0 \text{ and } y = 0. \end{cases}$$

3.7.



3.8. Up to 100, for example

$$f(x) = x + x(x-1)(x-2)(x-3)\dots(x-100).$$

As few as 0, for example

$$f(x) = x^{100} + x + 1.$$

3.9. 9, 4, 1, 1, 2

4.4. In one variable, you could try $f(x) = x^2$ or $f(x) = x^3$ as examples.

4.7. The kernel of $F'(x, \lambda, t)$ is given by the equations on (y, μ, s):

$$(A(t) - \lambda I)\, y = -A'(t)sx + \mu x, y_1 = 0, s = 0.$$

The only points where $F = 0$ are those with $t = 0, A(0)x = \lambda x, x_1 = 1$, so $t = 0$ and x is a λ-eigenvector. Take the particular choice of x and $\lambda = \lambda_0$ given in the problem. Look at the equations of the kernel of $F'(x, \lambda, t)$ at that (x, λ_0, t) point:

$$(A_0 - \lambda_0 I)\, y = \mu x, y_1 = 0, s = 0.$$

But then

$$(A_0 - \lambda_0 I)^2\, y = \mu\, (A_0 - \lambda_0 I)\, x = 0.$$

Since the generalized λ_0-eigenspace has dimension 1, y must be a λ_0-eigenvector. But again since the generalized λ_0-eigenspace has dimension 1, $y = cx$ for some c. But $y_1 = 0$ while $x_1 = 1$, so $c = 0$, i.e. $y = 0$. But then our equations become $0 = \mu x$ and $x \neq 0$ so $\mu = 0$. Therefore $(y, \mu, s) = (0, 0, 0)$, i.e $F'(x, \lambda, t)$ has kernel $(0, 0, 0)$, so is invertible. Therefore F is invertible near $(x, \lambda_0, 0)$, say with inverse $G(p, \sigma, t)$. We then let $(x(t), \lambda(t), t) = G(0, 0, t)$, and by definition of F, $F(x(t), \lambda(t), t) = (0, 0, t)$, i.e. $A(t)x(t) = \lambda(t)x(t)$ and $x_1(t) = 0$.

5.1.

$$F'(0,0,1) = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Since the y and z columns (i.e. second and third columns) are linearly independent, i.e. form an invertible matrix, the implicit function theorem says that we can solve for y and z as continuously differentiable functions of x near $(1, 0, 0)$.

The implicit function theorem doesn't tell us anything about solving for x and y as functions of z, since the x and y columns (the first and second) of $F'(0, 0, 1)$ are not linearly independent. Your drawing will show you that there are actually 2 choices of x and y with given z near $(x, y, z) = (1, 0, 0)$ so that $F(x, y, z) = 0$.

5.3. Let

$$F(x,y,z) = \begin{pmatrix} x^2 + y^2 - z \\ x^2 + z^2 - 1 \end{pmatrix}.$$

So C is the set where $F = 0$. Calculate

$$\frac{\partial F}{\partial(y,z)}(x,y,z) = \begin{pmatrix} 2y & -1 \\ 0 & 2z \end{pmatrix}.$$

This matrix is invertible except where $y = z = 0$. To be on the paraboloid and have $y = z = 0$ we must be at the origin. So the implicit function theorem says that near every point of C, except perhaps near the origin, the curve C can be somehow expressed as the graph $y = y(x), z = z(x)$ of continuously differentiable functions.

5.5.

$$\begin{pmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{pmatrix} = 2 \begin{pmatrix} y & z \\ y & 0 \end{pmatrix},$$

so has determinant $-yz$. Therefore the implicit function theorem applies except at the points where $-yz = 0$ and $x^2 + y^2 + z^2 = 1$ and $(x - a)^2 + y^2 = r^2$. So these are the points

1. $y = 0$ and $x = a \pm r$ and $z = \pm\sqrt{1 - x^2}$: the
2. $z = 0$ and $x^2 + y^2 = 1$ and $(x - a)^2 + y^2 = r^2$.

5.6. Clearly $F(Q, R) = 0$ just when Q is orthogonal and R is upper triangular. Pick some (Q_0, R_0) with

$$F(Q_0, R_0) = 0$$

and with R_0 invertible. Compute

$$F(Q_0 + Q_1, R_0 + R_1) - F(Q_0, R_0) = \left(Q_0 {Q_1}^t + Q_1 {Q_0}^t, L(R_1)\right) + \dots,$$

so that

$$F'(Q_0, R_0)(Q_1, R_1) = \left(Q_0 {Q_1}^t + Q_1 {Q_0}^t, L(R_1)\right).$$

It is easy to check that (Q_1, R_1) lies in the kernel just when $Q_1 = T_1 Q_0$ for a unique antisymmetric matrix T_1 and $L(R_1) = 0$ and $0 = Q_0 R_1 + T_1 Q_0 R_0$, so

$$R_1 = -{Q_0}^t T_1 Q_0 R_0,$$

which we write as

$$R_0^{-1} R_1 = -{Q_0}^t T_1 Q_0.$$

By the previous parts of the problem, the left hand side is upper triangular and that the right hand side is antisymmetric. Therefore both sides vanish.

5.7.

$$F'(L_0, U_0)(L_1, U_1) = L_0 U_1 + L_1 U_0.$$

The kernel is then given by the solutions to

$$L_0^{-1} L_1 = -U_1 U_0^{-1}.$$

The left hand side is lower triangular (why?) while the right hand side is upper triangular. So the kernel is 0. So F is locally invertible.

5.8. Try

$$F \colon \mathbb{R}^{n^2} \times \mathbb{R}^{n(n+1)/2} \to \mathbb{R}^{n(n+1)/2} \times \mathbb{R}^{n^2}, F(K, P) = \left(KK^t - I, KP\right),$$

so K is any matrix, and P is symmetric. Check that

$$F'(K, P)(L, Q) = \left(KL^t + LK^t, KQ + LP\right).$$

If we can show that $F'(K, P)$ has kernel $\{0\}$, then F is locally invertible by the inverse function theorem. But if $F'(K, P)(L, Q) = 0$, then

$$KL^t + LK^t = 0$$

so that LK^t is antisymmetric. If we work at a point where $K^t K = I$, then we easily see that LK^t is antisymmetric just when $K^t L$ is antisymmetric. Our second equation is $KQ + LP = 0$, which forces $QP^{-1} = -K^t L$, antisymmetric. This easily forces PQ to be antisymmetric. But now plug in that P is diagonal, without loss of generality, and assume that P has positive eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Then

$$PQ = (\lambda_i Q_{ij}) = -QP = (-\lambda_j Q_{ij}).$$

Adding, we get $(\lambda_i + \lambda_j)\, Q_{ij} = 0$, so that $Q_{ij} = 0$ for all i and j, i.e. $Q = 0$. But then $KQ + LP = 0$ so $LP = 0$ so $L = 0$. Therefore $F'(K, P)$ has kernel $\{0\}$. So F has a local continuously differentiable inverse, say F^{-1}, and we let $(K, P) = F^{-1}(0, A)$.

6.5. Looking at the equations, we see that on the intersection, $y \ge 0$ and $z \ge 0$. Let

$$F(x, y, z) = \left(x^2 + y^2 - z, x^2 + z^2 - y\right).$$

Then

$$F'(x, y, z) = \begin{pmatrix} 2x & 2y & -1 \\ 2x & -1 & 2z \end{pmatrix}.$$

Looking at 2×2 determinants, we can see that the matrix has full rank except at points where either $x = 0$ or $y = -1/2$ or $z = -1/2$. But $y \ge 0$ and $z \ge 0$, so only $x = 0$ can happen on the intersection. On $x = 0$, the rank drops only when $4yz = -1$. But $y \ge 0$ and $z \ge 0$ so $4yz \ge 0$. Therefore F' has rank 2 at every point of the intersection, so the intersection, being the zero locus of F, is a manifold. Since F' has rank 2, the set where $F = 0$ is a manifold of dimension $3 - 2 = 1$.

6.6.

$$F'(x, y, z) = 2 \begin{pmatrix} x & y & z \\ x - a & y & 0 \end{pmatrix}.$$

Suppose that $x^2 + y^2 + z^2 = 1$ and $(x - a)^2 + y^2 = r^2$. At any point of $x^2 + y^2 + z^2 = 1$, we find the first row is not zero. The implicit function theorem fails to apply just where F' is not onto, i.e. has rank less than 2, so just where the second row is a multiple of the first one, say

$$\begin{pmatrix} x - a & y & 0 \end{pmatrix} = \lambda \begin{pmatrix} x & y & z \end{pmatrix}$$

for some constant λ. If $\lambda = 0$, then $x = a$ and $y = 0$. But $(x - a)^2 + y^2 = r^2$, forcing $r = 0$, a contradiction. So $\lambda \ne 0$. But then

$$\begin{pmatrix} x - a & y & 0 \end{pmatrix} = \lambda \begin{pmatrix} x & y & z \end{pmatrix}$$

forces $z = 0$, and either $y = 0$ or $\lambda = 1$.

If $y = 0$, then $(x, y, z) = (\pm 1, 0, 0)$ and $a \pm r = \pm 1$. Clearly $a + r \ge 0$.

If $\lambda = 1$, then $x - a = x$ and $z = 0$, so $x^2 + y^2 = 1$ and $a = 0$ so $x^2 + y^2 = r^2$. So $r = 1$.

1. $(x, y, z) = (1, 0, 0), a + r = 1$, sphere intersecting cylinder in a “figure 8” shaped curve, continuously differentiable except at that one point where the sphere and cylinder are tangent.
2. $(x, y, z) = (1, 0, 0), a - r = 1$, sphere outside cylinder, touching at one point.
3. $(x, y, z) = (-1, 0, 0), a - r = -1$, sphere inside cylinder, touching at one point.
4. $(x, y, z) = (x, y, 0), a = 0, r = 1$, sphere inside cylinder, intersecting along the entire equator, which is continuously differentiable, but the implicit function theorem doesn’t know that.

6.7. Let

$$F(x, y, z) = \left(x^2 + y^2 - z, x^2 + z^2 - 4\right).$$

Then

$$F'(x, y, z) = \begin{pmatrix} 2x & 2y & -1 \\ 2x & 0 & 2z \end{pmatrix}.$$

Looking at 2×2 determinants, we can see that the matrix has full rank except at points where either $x = 0$ or $y = 0$. If $x = 0$, we still have full rank unless either $y = 0$ or $z = 0$. If $y = 0$, we still have full rank unless either $x = 0$ or $z = -\frac{1}{2}$. There are 3 cases of rank 1:

1. $x = 0$ and $y = 0$,
2. $x = 0$ and $z = 0$,
3. $y = 0, z = -\frac{1}{2}$.

The first is impossible at any point of the intersection since being on the paraboloid implies $z = 0$, but being on the cylinder implies $x^2 + z^2 = 4$. The second is impossible at any point of the intersection since being on the cylinder implies $x^2 + z^2 = 4$. The third is impossible at any point of the intersection since being on the paraboloid implies $z \geq 0$. Therefore F' has rank 2 at every point of the intersection, so the intersection, being the zero locus of F, is a manifold. Since F' has rank 2, the set where $F = 0$ is a manifold of dimension $3 - 2 = 1$.

6.8. Take chart $F(x) = (x, f(x))$, with left inverse $G(x, y) = x$.

6.9. Since the result is purely local, we can assume that M is the image of a single chart. Suppose that $(x(t), y(t)) = f(t)$ is such a chart, say with left inverse g. Then $g(x(t), y(t)) = t$. Differentiate to see that

$$g'(x(t), y(t)) \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = 1.$$

Therefore

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \neq 0.$$

So $x'(t) \neq 0$ or $y'(t) \neq 0$. Suppose that $x'(t) \neq 0$. By the inverse function theorem, there is a local inverse function, say $t = t(x)$. Since the result is local, we can assume that this is a global inverse function. So each point $(x, y) \in M$ is

$$(x, y) = (x(t), y(t)) = (x, y(t(x))).$$

The same idea works with any number of variables; it helps to permute the variables into order so that the derivative of the chart has an invertible square matrix in its top rows.

6.15. A square.

6.16. No: for example, let M be the image of the unit square under the map $(x, y) \mapsto (x + y, y)$: □ ▱

6.24.

$$h(R, \alpha, \beta) = (r, \theta, Z) = (R \cos \beta, \alpha, R \sin \beta).$$

You can compute h' anywhere, but it is easiest at $(1,0,0)$, where you find $h'(1,0,0)=I$. Since h is a diffeomorphism defined on a connected open set, $\det h'(R,\alpha,\beta)$ can't change sign, so must be positive everywhere.

6.26. If f is a chart for M and g is a chart for N, make a chart $F(x,y)=(f(x),g(y))$ for $M\times N$. The various charts obtained this way cover $M\times N$ and are compatibly oriented.

6.29. For $x\in\mathbb{R}$ let $F(x)=x^2$, so $M=\{0\}$, but $F'(0)=0$, so as a 1×1 matrix, $F'(0)$ has rank 0, *not* rank 1.

8.4. $F\left(x_1(t),x_2(t)\right)=(1,t+e)$.

8.10. For example, on $\mathbb{R}$, take

$$F(x)=\begin{cases} x^2 & \text{if } x\ge 0,\\ -x^2 & \text{if } x<0,\end{cases}$$

and $X(x)=1$.

9.1. Plug in the vector fields $e_1,e_2,\dots,e_p$ to see that $a_{12\dots p}=0$. Similarly for all of the other a_I.

9.3. $2\,dx\wedge dy\wedge du\wedge dv$

11.2. Along the bottom, $y=0$ constant so $dy=0$, and the same along the top: $y=1$ constant so $dy=0$.

$$\begin{aligned}\int_0^1 x\,dy\Big|_{y=0}+\int_0^1 1\,dy\Big|_{x=1}-\int_0^1 x\,dy\Big|_{y=1}-\int_0^1 0\,dy\Big|_{x=0}&=0+1-0-0,\\ &=1.\end{aligned}$$

11.3. The boxes X and Y have lower dimension than $X\times Y$; apply induction and you can reduce to orienting rectangles in the plane (easy) and intervals of the line (easy).

11.6. If $\omega=f(x)dx_1\wedge dx_2\wedge\dots\wedge dx_n$. then let

$$g(x)=\int_0^{x_1} f(x)dx_1$$

and let

$$\xi=g(x)\,dx_2\wedge dx_3\wedge\dots\wedge dx_n.$$

11.7. We can change ω to be 0 away from $F(X)\cup G(X)$, throughout $\mathbb{R}^n$. As we saw in problem 11.6 on page 75 $\omega=d\xi$ for some $(n-1)$-form $d\xi$, so

$$\begin{aligned}\int_X F^*\omega&=\int_{\partial X}F^*\xi,\\ &=\int_{\partial X}G^*\xi,\\ &=\int_X G^*\omega\end{aligned}$$

by Stokes's theorem in a box.

12.1. If we let

$$\omega = f(y)\, dy_1 \wedge dy_2 \wedge \cdots \wedge dy_n,$$

and let $y = F(x)$, then we get

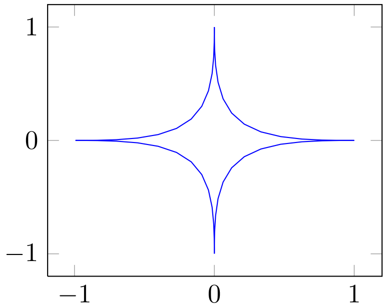
$$\begin{aligned}\int_{F(X)} f(y) &= \int_{F(X)} \omega, \\ &= \pm \int_X F^*\omega,\end{aligned}$$

with the sign $\pm$ depending on whether F is orientation preserving or reversing,

$$= \pm \int_X f(F(x)) \det F'(x) dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n.$$

The signs cancel: $\pm \det F'(x) = |\det F'(x)|$.

12.4. It helps to draw the curve:



Parameterize the boundary by

$$x(t) = \cos^5 t,\ \ y(t) = \sin^5 t$$

and compute $\int_\partial \omega$ where $\omega = \frac{1}{2}x\, dy - \frac{1}{2}y\, dx$.

12.6.

$$\begin{aligned}dx &= \sin\alpha\cos\beta\, dr + r\ \cos\alpha\cos\beta\, d\alpha - r\ \sin\alpha\sin\beta\, d\beta, \\ dy &= \sin\alpha\sin\beta\, dr + r\ \cos\alpha\sin\beta\, d\alpha + r\ \sin\alpha\cos\beta\, d\beta, \\ dz &= \cos\alpha\, dr - r\ \sin\alpha\, d\alpha.\end{aligned}$$

Wedging these together, a long calculation yields

$$dx \wedge dy \wedge dz = r^2 \sin\alpha dr \wedge d\alpha \wedge d\beta.$$

Let

$$Y = \{\, (r,\alpha,\beta) \mid 0 \le r \le 1, 0 \le \beta \le 2\pi, 0 \le \alpha \le \pi/2 \,\}.$$

The integral is

$$\int_X \left(1-\sqrt{x^2+y^2+z^2}\right)\,dx\wedge dy\wedge dz=\int_Y (1-r)\,r^2\,dr\sin\alpha d\alpha d\beta,$$
$$=\int_0^1 (1-r)r^2\,dr\int_0^{\pi/2}\sin\alpha\int_0^{2\pi}d\beta$$

and we let the reader fill in the remaining details.
12.7.

$$\begin{aligned}dx&=\cos\theta\,dr-r\,\sin\theta\,d\theta,\\ dy&=\sin\theta\,dr+r\,\cos\theta\,d\theta,\\ dz&=dz.\end{aligned}$$

i.e. $F^*dx=\cos\theta\,dr-r\,\sin\theta\,d\theta$, etc. Wedge these to find

$$dx\wedge dy\wedge dz=r^2\,dr\wedge d\theta\wedge dz.$$

and

$$F^{-1}X=\{\,(r,\theta,z)\mid 2\le r\le 3, 0\le z\le 1, \pi/4\le\theta\le\pi/2\,\}.$$

(You really have to draw it to find the correct θ interval.)

$$\int_X z\left(x^2+y^2\right)\,dx\wedge dy\wedge dz=\int_{r=2}^{r=3}\int_{\theta=\pi/4}^{\theta=\pi/2}\int_{z=0}^{z=1} zr^4\,dr\wedge d\theta\wedge dz,$$
$$=\int_2^3 r^4\,dr\frac{\pi}{4}\int z\,dz$$

and we leave the rest to the reader.
13.5.

$$\Omega=\frac{x}{r^3}\,dy\wedge dz+\frac{y}{r^3}\,dz\wedge dx+\frac{z}{r^3}\,dx\wedge dy.$$

13.6. $\Omega=\sin\alpha\,d\alpha\wedge d\beta$.
13.13. Define

$$K\colon \bar{B}\times\mathbb{R}\to S, K(x,t)=(1-t)x+t\,F(x).$$

Fix some $x\in\bar{B}$. Consider the function $q(t)=\|K(x,t)\|-1$. Expand out $q(t)$ to see exactly where t has to be so that $q(t)=0$, i.e. so that $K(x,t)\in S$. Show that $q(t)$ has one root ≥ 0 and one negative root, and that if we take the root $t=t(x)\ge 0$, then $t(x)$ is continuously differentiable for $x\in B$.

13.16. If X has no zero, then F is defined on $\bar{B}$ and continuously differentiable. On the sphere $S = \partial \bar{B}$, clearly $F(x) = x$ for all $x \in S$, so $F^* \Omega = \Omega$ and

$$\begin{aligned} 0 &< \int_S \Omega, \\ &= \int_S F^* \Omega, \\ &= \int_{\bar{B}} dF^* \Omega, \\ &= \int_{\bar{B}} F^* d\Omega, \\ &= 0, \end{aligned}$$

a contradiction, so X must have a zero.

15.4.

$$dV_{\text{ellipsoid}} = \frac{1}{\sqrt{\sum_i \left(\frac{x_i}{a_i}\right)^2}} \sum_j \frac{x_j}{a_j} dx_1 \wedge dx_2 \wedge \dots \wedge \widehat{dx_j} \wedge \dots \wedge dx_n.$$

16.2. Suppose that $f \colon P \to Q$ is a continuously differentiable map and that f' has constant rank. Then every level set $M = f^{-1}\{q_0\}$ is a manifold with tangent space

$$T_p M = \operatorname{kernel} f'(p)$$

for every point $p \in M$.

16.3. It looks impossible: you have to cover with disks, spaced out from one another so that they cover the line without "bunching up", but that won't allow the disks to have small radii, because each disk has to overlap with the next disk. However, if you let the radii be $r_j = 1/j$ and the centers be $x_1 = 0$ and $x_{j+1} = x_j + r_j/2$, for example, you can cover "half" the line, by the nonconvergence of the harmonic series, while the finiteness of the area sum of the disks is due to the convergence of $\sum 1/j^2$. See figure 16.1 on page 105. Then cover the "other half " in the same way. Rescale the picture to get the total area to be as small as you like.

16.4. If we can cover in balls with small volume sum, we can then cover each small ball by a small box, so clearly get small volume sum. So we need only worry about the other direction: if A is a set of measure zero using box coverings, is it also a set of measure zero using ball coverings? Take a box $X \subset \mathbb{R}^n$. We want to cover X by balls, so that the volume sum of the balls is close to $\operatorname{Vol}(X)$. We can increase the side lengths of X, making a bigger box with sides of rational lengths, say p_i/q_i, only slightly increasing the volume. By taking least common multiples, we can arrange that all of the p_i/q_i have the same denominator: $p_i/q_i = p_i/q$. Then we can chop X into a grid, with p_i subdivisions of side i of X, so into boxes whose every side has length $1/q$. Each of these small boxes fits into a ball of radius $r = \sqrt{n}/2q$. The volume of such a ball is

$$\operatorname{Vol}(B_r) = \operatorname{Vol}(B_1)\, r^n,$$

as we saw on page 87. Therefore the volume sum of all these balls is

$$p_1 p_2 \dots p_n \left(\frac{\sqrt{n}}{2q}\right)^n \operatorname{Vol}(B_1) = \left(\frac{\sqrt{n}}{2}\right)^n \operatorname{Vol}(B_1) \operatorname{Vol}(X).$$

So if we can cover a set by boxes with arbitrarily small volume sum, then we can cover it by balls so that we increase the volume sum by as close as we like to a constant factor of

$$\left(\frac{\sqrt{n}}{2}\right)^n \operatorname{Vol}(B_1),$$

so that we can still make the volume sum as small as we need.

16.5. From the solution of problem 16.4 on page 106, boxes, rotated or not, can be covered by balls, increasing the sum of volumes by only a constant factor. So if there is a covering by rotated boxes, with small volume sum, then there is a covering by balls with small volume sum.

16.8. Rough ideas: it is enough to prove the result for $f \circ h_1, f \circ h_2, \dots$, where these $h_1, h_2, \dots$ are a sequence of charts covering P. (Is there such a sequence? You decide.) So we can assume that $P \subset \mathbb{R}^p$ is an open subset. Covering P in a sequence of smaller open sets, each of which has image in the image of a single chart set in Q, we can similarly arrange that Q is an open set, say in $\mathbb{R}^{p+q}$. Let $F(s,t) = f(s)$ for $s \in \mathbb{R}^p$ and $t \in \mathbb{R}^q$, wherever this is defined. All points of the image of F (i.e. the image of f) are critical points. So they lie in a measure zero set by Sard's theorem.

17.2. A polynomial function of one variable is proper just when it isn't constant. In many variables, it is more complicated.

17.3. In order to get $f^{-1}U_Q$ to sit in U_P, we need U_Q to be contained in $\left(-\varepsilon^k, \varepsilon^k\right)$ if k is odd, and contained in $\left(-\infty, \varepsilon^k\right)$ if k is even.

Bibliography

[1] Vladimir I. Arnold, *Ordinary differential equations*, Universitext, Springer-Verlag, Berlin, 2006, Translated from the Russian by Roger Cooke, Second printing of the 1992 edition. MR 2242407 (2007b:34001) 53, 55, 115

[2] Raoul Bott and Loring W. Tu, *Differential forms in algebraic topology*, Graduate Texts in Mathematics, vol. 82, Springer-Verlag, New York, 1982. MR 658304 (83i:57016) iii

[3] Victor Guillemin and Alan Pollack, *Differential topology*, AMS Chelsea Publishing, Providence, RI, 2010, Reprint of the 1974 original. MR 2680546 (2011e:58001) iii

[4] John Hamal Hubbard and Barbara Burke Hubbard, *Vector calculus, linear algebra, and differential forms*, Prentice Hall Inc., Upper Saddle River, NJ, 1999, A unified approach. MR 1657732 (99k:00002) iii

[5] John W. Milnor, *Topology from the differentiable viewpoint*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, Based on notes by David W. Weaver, Revised reprint of the 1965 original. MR 1487640 (98h:57051) iii

[6] Walter Rudin, *Principles of mathematical analysis*, third ed., McGraw-Hill Book Co., New York, 1976, International Series in Pure and Applied Mathematics. MR 0385023 (52 #5893) iii

[7] Michael Spivak, *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*, W. A. Benjamin, Inc., New York-Amsterdam, 1965. MR 0209411 (35 #309) iii, 10, 110

[8] Shlomo Sternberg, *Lectures on differential geometry*, second ed., Chelsea Publishing Co., New York, 1983, With an appendix by Sternberg and Victor W. Guillemin. MR 891190 (88f:58001) 110

List of Notation

$\mathbb{R}^n$	the set of all n-tuples of real numbers	1
$\mathbb{R}$	the set of real numbers	1
$\mathbb{Z}$	the set of integers	1
$f(S)$	image of set S via the map f	1
$f^{-1}T$	preimage of set T via the map f	1
$B_r(x)$	open ball of radius r about the point x	2
$\bar{B}_r(x)$	closed ball of radius r abour the point x	2
$\langle x, y \rangle$	inner product of vectors x and y	2
$\|\|x\|\|$	norm of a vector x	2
$\sup_{x \in S} x$	supremum of a set S of real numbers	4
$\bar{A}$	closure of set A	5
$o(h)$	a quantity whose ratio to $\|\|h\|\|$ vanishes	11
$o(h)^2$	a quantity whose ratio to $\|\|h\|\|^2$ vanishes	16
A^t	transpose of a matrix A	36
∂M	boundary of a manifold M	42
$F_* X$	pushforward of a vector field	54
$\wedge$	wedge product of differential forms	60
$F^* \omega$	pullback of differential form	63
$d\theta$	angle 1-form	85
$\lrcorner$	hook	86
$T_x M$	tangent space to the manifold M at the point x	91
S^n	the n-dimensional unit sphere about 0 in $\mathbb{R}^{n+1}$	97
$\mathrm{Vol}(M)$	volume of a manifold M	100
$\deg_q f$	degree of a proper map f at a regular value q	111
$\deg f$	degree of a proper map	117

Index

affine map, ***71***
arc length form, ***103***
area, ***99***
area element, ***103***
area form, ***99***

ball
    closed, ***2***
    open, ***2***
bounded, ***2***
box, ***2***
bump function, ***15***

Cauchy integral theorem, 82
chart, ***39***
    positively oriented, ***44***
closed
    form, ***75***
closed set, ***3***
coconut theorem, ***93***
coffee, 89
compact, ***5***
continuous, ***3***
continuously differentiable, ***13***, 13
contraction, ***25***
contraction map, ***25***
convergence, ***4***
coordinates
    cylindrical, ***83***
    spherical, ***83***
corner, ***111***
critical points, ***19***
curve, ***46***
cylindrical coordinates, ***83***

degree, ***111***
diffeomorphic, ***39***
diffeomorphism, ***39***, 53
differentiable, ***11***
    continuously, ***13***, 13
differential, ***63***
differential form, ***58***, ***93***
differentiation, ***11***
dilation, ***25***
distance, ***2***
domain, ***81***, 83

exact
    form, ***75***

flow
    of vector field, ***116***
flow line, ***53***
form, ***58***
    closed, ***75***
    exact, ***75***
function, ***1***

Gauss map, ***119***

hairy ball theorem, ***95***
hedgehog theorem, ***95***
holomorphic, ***81***
homotopy, ***111***

image, ***1***
implicit function, ***33***
inner product, ***2***
integral, ***50***
    of form, ***69***
isotopy, ***116***

least upper bound, ***4***
left half space, ***42***
Leibnitz rule, ***65***
length element, ***103***
level set, ***33***
limit, ***4***
locally diffeomorphic, ***39***

Möbius strip, ***43***

manifold, ***39***
    with boundary, ***42***
    with boundary and corners, ***42***
map, ***1***
    contraction, ***25***
measure zero, ***105***

neat map, ***111***
neighborhood, ***3***
norm, ***2***
    of a matrix, ***25***

open cover, ***6***
open set, ***3***
orientation
    of a curve, ***43***
    of a manifold, ***44***
    of a surface, ***43***
orientation preserving, ***43***

paper, 89
partition of unity, ***8***
preimage, ***1***
proper, ***111***
pullback, ***63***
    form, ***63***
pushforward
    vector field, ***54***

rank theorem, ***35***, ***42***

saddle point, ***19***
Sard's theorem, ***108***
set
    closed, ***3***
    open, ***3***
sign
    of map between manifolds of equal dimension, ***111***
solid angle form, ***86***
spherical coordinates, ***83***
standard basis, ***2***
Stokes's theorem, ***80***
subcover, ***6***
support, ***5***
supremum, ***4***
surface, ***46***

tangent space, ***91***
tangent vector, ***91***
tangent vector field, ***92***
theorem
    Cauchy integral, 82
    coconut, ***93***
    contraction mapping, ***25***
    hairy ball, ***95***
    hedgehog, ***95***
    implicit function, ***33***
    rank, ***35***, ***42***
    Sard's, ***108***
    Stokes's, ***80***

vector field, ***53***
    tangent, ***92***
volume, ***99***
volume element, ***103***
volume form, ***99***

wedge product, ***60***